

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:33:07 ; Search time 34.09 Seconds

(without alignments)
550.660 Million cell updates/sec

Title: US-09-508-418-2

Sequence: 1 MTTTPINHNPIFHSSSS.....EGAYEVASEVTGELSRVAK 548

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 2821 | 99.2 | 548 | 1 | PROG_TOBAC |
| 2 | 2075 | 73.0 | 537 | 1 | PROG_ARATH |
| 3 | 361 | 15.0 | 471 | 1 | PROX_MYXA |
| 4 | 505 | 17.8 | 504 | 1 | PROX_MYXA |
| 5 | 485 | 17.1 | 470 | 1 | PROX_MYXA |
| 6 | 375.5 | 13.2 | 527 | 1 | PROX_MYXA |
| 7 | 362.5 | 12.8 | 477 | 1 | PROX_MYXA |
| 8 | 362.5 | 12.7 | 477 | 1 | PROX_MYXA |
| 9 | 302.5 | 10.6 | 490 | 1 | PROX_MYXA |
| 10 | 221 | 7.8 | 426 | 1 | PROX_MYXA |
| 11 | 216.5 | 7.6 | 451 | 1 | PROX_MYXA |
| 12 | 189 | 6.6 | 539 | 1 | PROX_MYXA |
| 13 | 141 | 5.0 | 448 | 1 | PROX_MYXA |
| 14 | 139 | 4.9 | 570 | 1 | PROX_MYXA |
| 15 | 134 | 4.7 | 522 | 1 | PROX_MYXA |
| 16 | 132.5 | 4.7 | 522 | 1 | PROX_MYXA |
| 17 | 131 | 4.6 | 566 | 1 | PROX_MYXA |
| 18 | 125.5 | 4.4 | 417 | 1 | PROX_MYXA |
| 19 | 125.5 | 4.4 | 472 | 1 | PROX_MYXA |
| 20 | 125.5 | 4.4 | 489 | 1 | PROX_MYXA |
| 21 | 123.5 | 4.3 | 526 | 1 | PROX_MYXA |
| 22 | 123 | 4.3 | 582 | 1 | PROX_MYXA |
| 23 | 122 | 4.3 | 471 | 1 | PROX_MYXA |
| 24 | 119.5 | 4.2 | 478 | 1 | PROX_MYXA |
| 25 | 118 | 4.1 | 571 | 1 | PROX_MYXA |
| 26 | 117.5 | 4.1 | 469 | 1 | PROX_MYXA |
| 27 | 116 | 4.1 | 583 | 1 | PROX_MYXA |
| 28 | 114.5 | 4.0 | 1106 | 1 | PROX_MYXA |
| 29 | 114 | 4.0 | 531 | 1 | PROX_MYXA |
| 30 | 112.5 | 4.0 | 570 | 1 | PROX_MYXA |
| 31 | 111 | 3.9 | 508 | 1 | PROX_MYXA |
| 32 | 109.5 | 3.9 | 527 | 1 | PROX_MYXA |
| 33 | 109.5 | 3.9 | 908 | 1 | PROX_MYXA |

| | | | | | | | |
|----|-------|-----|------|---|------------|--------|-------------|
| 34 | 109 | 3.8 | 1181 | 1 | ITR7_HUMAN | Q13683 | homo sapien |
| 35 | 108.5 | 3.8 | 1062 | 1 | CERU_MOUSE | Q61147 | mus musc |
| 36 | 108 | 3.8 | 504 | 1 | A37C_DROME | P18487 | drosophila |
| 37 | 107.5 | 3.8 | 516 | 1 | ER13_ARATH | Q65403 | arabidopsi |
| 38 | 107 | 3.8 | 405 | 1 | SAOX_CORS1 | P40875 | corynebacte |
| 39 | 107 | 3.8 | 1179 | 1 | ITR7_MOUSE | Q61738 | mus muscu |
| 40 | 106.5 | 3.7 | 630 | 1 | ITR1_MOUSE | Q09046 | mus muscu |
| 41 | 105.5 | 3.7 | 872 | 1 | IP3L_RAT | P42335 | rattus norv |
| 42 | 105 | 3.7 | 765 | 1 | TR2M_AGR3 | P25017 | agrobacteri |
| 43 | 104.5 | 3.7 | 466 | 1 | AMID_PSEPU | Q69768 | pseudomonas |
| 44 | 104.5 | 3.7 | 517 | 1 | ER12_ARATH | Q65402 | arabidopsi |
| 45 | 104.5 | 3.7 | 524 | 1 | CRT1_RHOCA | P17054 | rhodobacter |

ALIGNMENTS

| RESULT | ID | PROG_TOBAC | STANDARD | PRT | 548 AA |
|--------|--|--------------------------------|----------|-----|--------|
| AC | 024163 | 15-DEC-1998 (Rel. 37, Created) | | | |
| DT | 15-DEC-1998 (Rel. 37, Last sequence update) | | | | |
| DT | 15-JUL-1999 (Rel. 38, Last annotation update) | | | | |
| DE | PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I) | | | | |
| DE | (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I). | | | | |
| GN | PPX1. | | | | |
| OS | Nicotiana tabacum (Common tobacco). | | | | |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; | | | | |
| OC | Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; | | | | |
| OC | Solanales; Solanaceae; Nicotiana. | | | | |
| OX | NCBI_TaxID=4097; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN=CV. SRL; | | | | |
| RX | MEDLINE=97385200; PubMed=9238074; | | | | |
| RA | Lermontova I., Kruse E., Mock H.-P., Grimm B.; | | | | |
| RT | "Cloning and characterization of a plastid and a mitochondrial | | | | |
| RT | isoform of tobacco protoporphyrinogen IX oxidase."; | | | | |
| RT | Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997). | | | | |
| CC | - FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN | | | | |
| CC | IX TO FORM PROTOPORPHYRIN IX. | | | | |
| CC | - CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN- | | | | |
| CC | IX + H(2)O(2). | | | | |
| CC | - COFACTOR: FAD (BY SIMILARITY). | | | | |
| CC | - PATHWAY: PENTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS. AND | | | | |
| CC | CHLOROPHYLL SYNTHESIS. | | | | |
| CC | - SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY | | | | |
| CC | SIMILARITY). | | | | |
| CC | - SUBCELLULAR LOCATION: CHLOROPLAST. | | | | |
| CC | - DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREATURE LEAVES. | | | | |
| CC | DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL | | | | |
| CC | IN ROOTS. | | | | |
| CC | - INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL | | | | |
| CC | EXPRESSION IN THE DARK PERIOD. | | | | |
| CC | - SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY. | | | | |
| CC | - This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| DR | EMBL: Y13465; CAA73865.1; | | | | |
| KM | porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase; | | | | |
| KM | Transit peptide; Flavoprotein; FAD; Chloroplast; | | | | |
| KM | Chlorophyll biosynthesis. | | | | |
| FT | TRANSIT 1 50 | | | | |
| FT | CHAIN 51 548 | | | | |
| FT | NP_BIND 78 83 | | | | |
| FT | SEQUENCE 548 AA: 59230 MM: 668922E78FB8A3E30 CR64; | | | | |

Query Match 99.2% Score 2821; DB 1: Length 548;
 Best Local Similarity 99.3% Pred. No. 14e-205;
 Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MTTPIANHPNFTTHOSSSSPLAFNLRTSIFPSSISKNSVNCNMRTRCSVAKDYVP 60
   |||||||
DB 1 MTTPIANHPNFTTHOSSSSPLAFNLRTSIFPSSISKNSVNCNMRTRCSVAKDYVP 60
QY 61 SSADVGGPAALDCVYVAGISGLCIAQVMSANTPMLVTEARBRAGNNTTVERDGYLM 120
   |||||||
DB 61 SSADVGGPAALDCVYVAGISGLCIAQVMSANTPMLVTEARBRAGNNTTVERDGYLM 120
QY 121 EGGPNSFOPSPDMLTMAVDCGLDLDVLDGPNAPRFVLMKGLRVPKIDLPEDDMS 180
   |||||||
DB 121 EGGPNSFOPSPDMLTMAVDCGLDLDVLDGPNAPRFVLMKGLRVPKIDLPEDDMS 180
QY 181 IPGKLRAFGPIGLRPSPPGHESEVBOFVRNLGGEVERLIEPPSGVYVGDPSKLSMK 240
   |||||||
DB 181 IPGKLRAFGPIGLRPSPPGHESEVBOFVRNLGGEVERLIEPPSGVYVGDPSKLSMK 240
QY 241 AAFGRVWKLBTGGSTIGGTFKAIKRSSTPRAPRDPRLPKKQGTGVSFRKGLRMLPDA 300
   |||||||
DB 241 AAFGRVWKLBTGGSTIGGTFKAIKRSSTPRAPRDPRLPKKQGTGVSFRKGLRMLPDA 300
QY 301 ISARLGSKLKLSMKLSSTKSEKGYHLYETPEGVVSLOSRSIYMTVPYVANSILRPL 360
   |||||||
DB 301 ISARLGSKLKLSMKLSSTKSEKGYHLYETPEGVVSLOSRSIYMTVPYVANSILRPL 360
QY 361 SVAAADALSNFYPPVGAIVTISTPOAIDERLVDELKKGCOLHPRTQVETIGTIYSS 420
   |||||||
DB 361 SVAAADALSNFYPPVGAIVTISTPOAIDERLVDELKKGCOLHPRTQVETIGTIYSS 420
QY 421 SLFPRNAPKRGVLLNVTGAKNPETLSTESOLVEVNDRIKMLIKPKODPLVGVVR 480
   |||||||
DB 421 SLFPRNAPKRGVLLNVTGAKNPETLSTESOLVEVNDRIKMLIKPKODPLVGVVR 480
QY 481 VMPQAIPOPLVGHLDLTSTAKAAMNDNGLEGLFLOGNVYSGVALRCVEGAYEVAEVTG 540
   |||||||
DB 481 VMPQAIPOPLVGHLDLTSTAKAAMNDNGLEGLFLOGNVYSGVALRCVEGAYEVAEVTG 540
QY 541 FLSTRAYK 548
   |||||||
DB 541 FLSTRAYK 548

```

RESULT 2

PROC_ARATH STANDARD: PRT: 537 AA.
 AC P55826;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO).
 GN PPOX OR AT4G01690 OR T15B16.13.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97136707; PubMed=8982084;
 RA Naita S., Tanaka R., Ito T., Okada K., Takekuni S., Inokuchi H.;
 RT "Molecular cloning and characterization of a cDNA that encodes
 RT protoporphyrinogen oxidase of Arabidopsis thaliana.";
 RL Gene 182:169-175(1996).
 [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CV, COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;
 RT "K.F.X., Schueller C., Wandutt R., Murphy G., Volckaert G.,

Query Match 99.2% Score 2821; DB 1: Length 548;
 Best Local Similarity 99.3% Pred. No. 14e-205;
 Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

RA Pohl T., Duesterhoeft A., Stiekema W., Ertan K.-D., Terryn N.,
RA Harris B., Anstorge W., Brandt P., Grivell L., Rieger M.,
RA Weichselgartner M., de Simone V., Obermayer B., Macho R., Mueller M.,
RA Krels M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hohnel S., Zimmermann W., Medler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Meljens I., Voet M., Bastiaens I., Aert R., Delcor E.,
RA Weizenecker T., Boche G., Kampsperger U., Hilbert H., Braun M.,
RA Holzner E., Brandt A., Peters S., Van Staveren M., Dirkse W.,
RA Moijman P., Klein Lanckhorst R., Rose M., Hauf J., Koester P.,
RA Berneiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McLay K., Hayes R.,
RA Petrett A., Rajandream M.-A., Lyne M., Benes V., Reclmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Partmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argitidou A., Vitale D., Liqiori R., Piravandi E.,
RA Massenet O., Oubley F., Clabaud G., Muedeln A., Reiber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chetdor F., Cooke R., Berger C., Montfort A., Casacuberta E.,
RA Glabons T., Weber N., Vandenbol M., Baryges M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Belke C.,
RA Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Beyan M., Wilson R.K., de la Bastide M., Haberman K.,
RA Parnell L., Dedina N., Guoj L., Schut K., Huang E., Spiegel L.,
RA Senkon M., Murray J., Sheet P., Cordes K., Abu-Thelidh J.,
RA Stoneking T., Kallio J., Graves T., Harmon G., Edwards J.,
RA Lacroille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Mix P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
RA Nelson J., Spleth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali T., Berghoff A., Jones K., Dione K., Cotton M., Joshi C.,
RA Anonolu B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vill D., Sheher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shobhy N., Hasegawa A., Hammed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana."
RT Nature 402:769-777(1999).
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED AT HIGH LEVELS IN THE LEAVES AND AT
CC LOW LEVELS IN THE ROOTS AND FLORAL BUDS.
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D83139; BAA11820.1;
CC DR EMBL: AL61492; CAB77739.1;
CC KW Protoporphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
CC Flavo-protein; FAD; Chloroplast; Transist peptide;
CC TRANSIT 1 7 CHLOROPLAST (POTENTIAL).
CC CHAIN ? 537 PROTOPORPHYRINOGEN OXIDASE.
CC NP_BIND 63 68 FAD (POTENTIAL).
CC SEQUENCE 537 AA; 57695 MW; C125B0EEIDJ3187D5 CRC64;

```


Thu Jul 5 14:36:07 2001

us-09-508-418-2.ra1

Page 1

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:01 ; Search time 28.95 Seconds

(Without alignments)
381.322 Million cell updates/sec

Title: US-09-508-418-2

Perfect score: 2844
Sequence: 1 MTTTPINHRIRFHSSSS.....EGAYEAVSEVTGLSRIRAK 548

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144636 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCNUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|--------|---------------|--------|----|--|
| 1 | 2113 | 74.3 | 539 | 2 | US-08-808-931-16 Sequence 16, Appl |
| 2 | 2113 | 74.3 | 539 | 3 | US-08-808-931-16 Sequence 16, Appl |
| 3 | 2113 | 74.3 | 539 | 3 | US-09-050-603A-16 Sequence 16, Appl |
| 4 | 2113 | 74.3 | 539 | 3 | US-09-102-420B-16 Sequence 16, Appl |
| 5 | 2075 | 73.0 | 537 | 1 | US-08-472-028A-2 Sequence 2, Appl |
| 6 | 2075 | 73.0 | 537 | 2 | US-08-808-931-2 Sequence 2, Appl |
| 7 | 2075 | 73.0 | 537 | 3 | US-08-808-931-2 Sequence 2, Appl |
| 8 | 2075 | 73.0 | 537 | 3 | US-09-050-603A-2 Sequence 2, Appl |
| 9 | 2075 | 73.0 | 537 | 3 | US-09-102-420B-2 Sequence 2, Appl |
| 10 | 2075 | 73.0 | 537 | 4 | US-09-071-296-2 Sequence 2, Appl |
| 11 | 2055 | 72.3 | 543 | 2 | US-08-808-931-12 Sequence 12, Appl |
| 12 | 2055 | 72.3 | 543 | 3 | US-08-808-931-12 Sequence 12, Appl |
| 13 | 2055 | 72.3 | 543 | 3 | US-09-050-603A-12 Sequence 12, Appl |
| 14 | 2055 | 72.3 | 543 | 3 | US-09-102-420B-12 Sequence 12, Appl |
| 15 | 2006.5 | 70.6 | 536 | 2 | US-08-808-931-20 Sequence 20, Appl |
| 16 | 2006.5 | 70.6 | 536 | 3 | US-08-808-931-20 Sequence 20, Appl |
| 17 | 2006.5 | 70.6 | 536 | 3 | US-09-050-603A-20 Sequence 20, Appl |
| 18 | 2006.5 | 70.6 | 536 | 3 | US-09-102-420B-20 Sequence 20, Appl |
| 19 | 1987 | 69.9 | 528 | 2 | US-08-808-931-10 Sequence 10, Appl |
| 20 | 1987 | 69.9 | 528 | 3 | US-08-808-931-10 Sequence 10, Appl |
| 21 | 1987 | 69.9 | 528 | 3 | US-09-050-603A-10 Sequence 10, Appl |
| 22 | 1987 | 69.9 | 528 | 3 | US-09-102-420B-10 Sequence 10, Appl |
| 23 | 1978 | 69.5 | 481 | 1 | US-08-472-028A-6 Sequence 6, Appl |
| 24 | 1978 | 69.5 | 481 | 2 | US-08-808-931-6 Sequence 6, Appl |
| 25 | 1978 | 69.5 | 481 | 3 | US-08-808-931-6 Sequence 6, Appl |
| 26 | 1978 | 69.5 | 481 | 3 | US-09-050-603A-6 Sequence 6, Appl |
| 27 | 1978 | 69.5 | 481 | 3 | US-09-102-420B-6 Sequence 6, Appl |

| | | | | | |
|----|--------|------|-----|---|--|
| 28 | 1978 | 69.5 | 483 | 4 | US-09-071-296-6 Sequence 6, Appl |
| 29 | 1933.5 | 68.0 | 560 | 2 | US-08-808-931-18 Sequence 18, Appl |
| 30 | 1933.5 | 68.0 | 560 | 3 | US-08-808-931-18 Sequence 18, Appl |
| 31 | 1933.5 | 68.0 | 560 | 3 | US-09-050-603A-18 Sequence 18, Appl |
| 32 | 1933.5 | 68.0 | 560 | 3 | US-09-102-420B-18 Sequence 18, Appl |
| 33 | 1823 | 64.1 | 440 | 2 | US-08-808-931-24 Sequence 24, Appl |
| 34 | 1823 | 64.1 | 440 | 3 | US-08-808-931-24 Sequence 24, Appl |
| 35 | 1823 | 64.1 | 440 | 3 | US-09-050-603A-24 Sequence 24, Appl |
| 36 | 1823 | 64.1 | 440 | 3 | US-09-102-420B-24 Sequence 24, Appl |
| 37 | 1261 | 44.3 | 312 | 2 | US-08-808-931-22 Sequence 22, Appl |
| 38 | 1261 | 44.3 | 312 | 3 | US-08-808-931-22 Sequence 22, Appl |
| 39 | 1261 | 44.3 | 312 | 3 | US-09-050-603A-22 Sequence 22, Appl |
| 40 | 1261 | 44.3 | 312 | 3 | US-09-102-420B-22 Sequence 22, Appl |
| 41 | 490.5 | 17.2 | 508 | 1 | US-08-472-028A-4 Sequence 4, Appl |
| 42 | 490.5 | 17.2 | 508 | 2 | US-08-808-931-4 Sequence 4, Appl |
| 43 | 490.5 | 17.2 | 508 | 3 | US-08-808-931-4 Sequence 4, Appl |
| 44 | 490.5 | 17.2 | 508 | 3 | US-09-050-603A-4 Sequence 4, Appl |
| 45 | 490.5 | 17.2 | 508 | 3 | US-09-102-420B-4 Sequence 4, Appl |

ALIGNMENTS

RESULT 1:
US-08-808-931-16
Sequence 16, Application US/08808931
Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Voltrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5939602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8687
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:00:01 ; Search time 28.95 Seconds
(without alignments)
381.322 Million cell updates/sec

Title: US-09-508-418-2
Perfect score: 2844
Sequence: 1 MTTPIANHPNIFTHOSSS.....EGAYEVASEVTEGLSRYAK 548

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA: *
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/1/1aa/PCRTS.COMB.pep: *
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|-------------------|
| 1 | 2113 | 74.3 | 539 | 2 | US-08-808-931-16 |
| 2 | 2113 | 74.3 | 539 | 3 | US-08-808-323-16 |
| 3 | 2113 | 74.3 | 539 | 3 | US-09-050-603A-16 |
| 4 | 2113 | 74.3 | 539 | 3 | US-09-102-420B-16 |
| 5 | 2075 | 73.0 | 537 | 1 | US-08-472-028A-2 |
| 6 | 2075 | 73.0 | 537 | 2 | US-08-808-931-2 |
| 7 | 2075 | 73.0 | 537 | 3 | US-08-808-323-2 |
| 8 | 2075 | 73.0 | 537 | 3 | US-09-050-603A-2 |
| 9 | 2075 | 73.0 | 537 | 3 | US-09-102-420B-2 |
| 10 | 2075 | 73.0 | 537 | 4 | US-09-071-296-2 |
| 11 | 2055 | 72.3 | 543 | 2 | US-08-808-931-12 |
| 12 | 2055 | 72.3 | 543 | 3 | US-08-808-323-12 |
| 13 | 2055 | 72.3 | 543 | 3 | US-09-050-603A-12 |
| 14 | 2055 | 72.3 | 543 | 3 | US-09-102-420B-12 |
| 15 | 2006.5 | 70.6 | 536 | 2 | US-08-808-931-20 |
| 16 | 2006.5 | 70.6 | 536 | 3 | US-08-808-323-20 |
| 17 | 2006.5 | 70.6 | 536 | 3 | US-09-050-603A-20 |
| 18 | 2006.5 | 70.6 | 536 | 3 | US-09-102-420B-20 |
| 19 | 1987 | 69.9 | 528 | 2 | US-08-808-931-10 |
| 20 | 1987 | 69.9 | 528 | 3 | US-08-808-323-10 |
| 21 | 1987 | 69.9 | 528 | 3 | US-09-050-603A-10 |
| 22 | 1987 | 69.9 | 528 | 3 | US-09-102-420B-10 |
| 23 | 1978 | 69.5 | 481 | 1 | US-08-472-028A-6 |
| 24 | 1978 | 69.5 | 481 | 2 | US-08-808-931-6 |
| 25 | 1978 | 69.5 | 481 | 3 | US-08-808-323-6 |
| 26 | 1978 | 69.5 | 481 | 3 | US-09-050-603A-6 |
| 27 | 1978 | 69.5 | 481 | 3 | US-09-102-420B-6 |

| | | | | | | |
|----|--------|------|-----|---|-------------------|-------------------|
| 28 | 1978 | 69.5 | 483 | 4 | US-09-071-296-6 | Sequence 6, Appl1 |
| 29 | 1933.5 | 68.0 | 560 | 2 | US-08-808-931-18 | Sequence 18, Appl |
| 30 | 1933.5 | 68.0 | 560 | 3 | US-08-808-323-18 | Sequence 18, Appl |
| 31 | 1933.5 | 68.0 | 560 | 3 | US-09-050-603A-18 | Sequence 18, Appl |
| 32 | 1933.5 | 68.0 | 560 | 3 | US-09-102-420B-18 | Sequence 18, Appl |
| 33 | 1823 | 64.1 | 440 | 2 | US-08-808-931-24 | Sequence 24, Appl |
| 34 | 1823 | 64.1 | 440 | 3 | US-08-808-323-24 | Sequence 24, Appl |
| 35 | 1823 | 64.1 | 440 | 3 | US-09-050-603A-24 | Sequence 24, Appl |
| 36 | 1823 | 64.1 | 440 | 3 | US-09-102-420B-24 | Sequence 24, Appl |
| 37 | 1261 | 44.3 | 312 | 2 | US-08-808-931-22 | Sequence 22, Appl |
| 38 | 1261 | 44.3 | 312 | 3 | US-08-808-323-22 | Sequence 22, Appl |
| 39 | 1261 | 44.3 | 312 | 3 | US-09-050-603A-22 | Sequence 22, Appl |
| 40 | 1261 | 44.3 | 312 | 3 | US-09-102-420B-22 | Sequence 22, Appl |
| 41 | 490.5 | 17.2 | 508 | 1 | US-08-472-028A-4 | Sequence 4, Appl1 |
| 42 | 490.5 | 17.2 | 508 | 2 | US-08-808-931-4 | Sequence 4, Appl1 |
| 43 | 490.5 | 17.2 | 508 | 3 | US-08-808-323-4 | Sequence 4, Appl1 |
| 44 | 490.5 | 17.2 | 508 | 3 | US-09-050-603A-4 | Sequence 4, Appl1 |
| 45 | 490.5 | 17.2 | 508 | 3 | US-09-102-420B-4 | Sequence 4, Appl1 |

ALIGNMENTS

RESULT 1
US-08-808-931-16
; Sequence 16, Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Voltrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Polyporphyrinogen Oxidase and Inhibitor-Resistant Mutants
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 539 amino acids
; TYPE: amino acid

APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
Protoporhyrinogen Oxidase
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6023012artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/050,603A
FILING DATE: 30-MAR-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-050-603A-16

Query Match 74.3%, Score 2113; DB 3; Length 539;
Best Local Similarity 79.2%, Pred. No. 9,4e-202;
Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;

QY 47 WTRCSVAKDITVPSAVDGGPAALDCYIVAGISGLCIAQVMSANT----PMLMATEA 102
DB 34 FLKRSIAEGPIITSSKIDGESSIADCVIVGGISGLCIAQALATKRDVAVSNVITREA 93
QY 103 RDRAGNITTVRDCGYLMEEGNSFQSPDPMUTJAVDGLDYLGLGNAPRVLWKGK 162
DB 94 RDRVGNITTVRDCGYLMEEGNSFQSPDPMUTJAVDGLDYLGLGNAPRVLWKGK 153
QY 163 LRVPVSKTLDLPFDLMSIPKLAGFGPIGLRSPPGHEBSVQFVRNNGEVFERLI 222
DB 154 LRVPVSKPTDLPFDLMSIAGKLAGFGAIGIRPPPGYEESEVFRNNGAVFERFI 213
QY 223 EPPCGYVVGSPKSLMSAAFGYKWKLEETGSGITGCFKAIKRSSSTPKAPRDPRLPKP 282
DB 214 EPPCGYVVGSPKSLMSAAFGYKWKLEETGSGITGCFKAIKRSSSTPKAPRDPRLPKP 273

QY 283 KQOTVGFPRKGLRMLPDALISARLSKLTSMKLSITSEKGGHLLTETPEGVYSIQR 342
DB 274 KQOTVGFPRKGLTLPDEALANSLSNVKLSITSLGNGVNLTFETPEGVYSIQR 333
QY 343 STVMTVPSYVASNILRPLSVAAADALSNFYYPYGVAVTISYPOAIRDERIVDELKFG 402
DB 334 SVMTIPSHVASNILHPLISAAADALSOFFYPPVAVSYVSPKFAIRKRECLIDELKFG 393
QY 403 QHPRPGVEYETIGTYSSELPNNRPAKGRVLLNLTGAKNPELISKTESQVEYVDRDL 462
DB 394 QHPRSGIETLGTYSSELPNNRPAKGRVLLNLTGAKNPELISKTESQVEYVDRDL 453
QY 463 RKLIRKPAODEPLVGVVWVMPAIPQFLVGHJDTLSTAKAMNNGELGFLGNVYSGV 522
DB 454 RKLMLNPNKODEPLVGVVWVMPAIPQFLVGHJDLSDAKMALRDSGRHGLFLGNTVSGV 513
QY 523 ALGRCEGAYEVAEYVGTSLRYAYK 548
DB 514 ALGRCEGAYEVAEYKELSOYAYK 539

RESULT 4
US-09-102-420B-16
Sequence 16, Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
OXIDASE ("PROTOX")
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995

```

ATTORNEY/AGENT INFORMATION:
NAME: Meliss J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEO ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 539 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-102-420B-16

```

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 74.38; | Score 2113; | DB 3; | Length 539; |
| Best Local Similarity | 79.28; | Pred. No. 9.4e-202; | | |
| Matches 401; | Conservative 46; | Mismatches 55; | Indels 4; | Gaps 1 |

[illegible]

```

: RESULT 5
: US-08-472-028A-2
: Sequence 2, Application US/08472028A
: Patent No. 5767373
:
: GENERAL INFORMATION:
: APPLICANT: Ward, Eric R
: APPLICANT: Volrath, Sandra
: TITLE OF INVENTION: Manipulation of Protoporphyrinogen
: TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
: NUMBER OF SEQUENCES: 12
:
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 7 Skyline Drive
: CITY: Hawthorne
: STATE: NY
:

```

```

1      COUNTRY: USA
2      ZIP: 10532
3      COMPUTER READABLE FORM:
4      MEDIUM TYPE: Floppy disk
5      COMPUTER: IBM PC compatible
6      OPERATING SYSTEM: PC-DOS/MS-DOS
7      SOFTWARE: PatentIn Release #1.0, Version #1.25
8      CURRENT APPLICATION DATA:
9      APPLICATION NUMBER: US/08/472,028A
10     FILING DATE:
11     CLASSIFICATION: 800
12     ATTORNEY/AGENT INFORMATION:
13     NAME: Elmer, James Scott
14     REGISTRATION NUMBER: 36,129
15     REFERENCE/POCKET NUMBER: CGC 1748/CIP
16     TELECOMMUNICATION INFORMATION:
17     TELEPHONE: 919-541-8614
18     TELEFAX: 919-541-8689
19     INFORMATION FOR SEQ ID NO: 2:
20     SEQUENCE CHARACTERISTICS:
21     LENGTH: 537 amino acids
22     TYPE: amino acid
23     TOPOLOGY: linear
24     MOLECULE TYPE: protein
25     US-08-472-028A-2

```

| | | | | |
|-----------------------|------------------|---------------------|------------|-------------|
| Query Match | 73.0%; | Score 2075; | DB 1; | length 537; |
| Best Local Similarity | 75.4%; | Pred. No. 5.7e-198; | | |
| Matches 404; | Conservative 49; | Mismatches 73; | Indels 10; | Gaps 4; |

| | | | | | |
|----|-----|--------------------------------|------------------------------|------------------------|-----|
| QY | 22 | IAFIKRTSTPTEPSISKSNVNCNGM--- | RIKCSVADIVFPSSAVUGFAMLE--- | DCVL | 70 |
| Db | 3 | LSLAPPTQSLSPFSKPN-LRLVYKFLRLRC | SVAGPPTVGSKIEGGGTTITTDVLI | 61 | |
| QY | 77 | VWAGISGLCIAQWMSAM----- | PNIMWTEARDRAGGNTTVERGYLMEGPN | SFOPS | 132 |
| Db | 62 | VGGISGLICIAQALTKRPDPA | PNLITTEADRVGNTITRENGFLMEGPN | SFOPS | 121 |
| QY | 133 | MLTMAVDCGLKDDVLGDPNAP | REFLWMLKGLRPVPSKLTLDLPEFDLMS | DPKLRAGEGP | 192 |
| Db | 122 | MLTMAVDCGLKDDVLGDPNAP | REFLWMLKGLRPVPSKLTLDLPEFDLMS | ISIGKIRAGGAL | 181 |
| QY | 193 | GLRPSPPGHEEVEBOFVVRN | NLGGEVEEFLIEPFGSGVYGDSPSKL | SKMAFGKWKLEET | 252 |
| Db | 182 | GRPSPPGHEEVEEVRNRNLG | DEVEERLIEPFGSGVYAGPDKLS | SKMAFGKWKLEON | 241 |
| QY | 253 | GGSIIGTFRKAIKERSSTP | KAPRDLRLPKPGQTVGSFRKGLRML | PDATISARLSKLS | 312 |
| Db | 242 | GGSIIGTFRKAIQERKNAP | KAEKRDRLPKPGQTVGSFRKGLRML | PEATISARLSKLS | 301 |
| QY | 313 | WKLSITKSEKGCYHLYTET | PEPCVVISLOSRSITVMTVP | SVYASNILRLPLSVAAALSNFY | 372 |
| Db | 302 | WKLSITKSEKGCYHLYTET | PEPCVVISLOSRSITVMTVP | SVYASNILRLPLSVAAALSNFY | 361 |
| QY | 373 | YRPVGVNTISYQOEAIR | DBRLVDGELKFGOLHPRTOGEV | ELGTYSSSLPEPNRPAKGRV | 432 |
| Db | 362 | YRPVGVNTISYQOEAIR | DBRLVDGELKFGOLHPRTOGEV | ELGTYSSSLPEPNRPAKGRV | 421 |
| QY | 433 | LTLNITGAKNPEILSKTES | QULEVEYVDDRLKMLIKP | KAODPLVVGARVWFOALPQRLV | 492 |
| Db | 422 | LTLNITGASTNGLISKSE | GELVEAVDDRLKMLIKP | STPDKLGIYVWFOALPQRLV | 481 |
| QY | 493 | HLDTLSAKAANNONGLE | BELEFGGVYGVGVALGR | VEGAEVASEVTEGELSLRAYK | 548 |
| Db | 482 | HLDTLSAKSITSSGVE | BELEFGGVYGVGVALGR | VEGAEVASEVTEGELSLRAYK | 537 |

RESULT 6
US-08-808-931-2
; Sequence 2, Application US/088089311
; Patent No. 5939602
; GENERAL INFORMATION:

APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
TITLE OF INVENTION: thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESS: No. 593602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-931-2

Query Match 73.0%; Score 2075; DB 2; Length 537;
Best Local Similarity 75.4%; Pred. No. 5.7e-198;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;
QY 22 LAFNRSTFIFSSISKSNVNCWG--RTRGVADYTPSSAVNGGPAEL-DCVI 76
DB 3 LSLRPTQSLPFSKFN-LRLNVYPLRLGCVAGGPVIGSSKIGGGTTITTCVI 61
QY 77 VAGISIGICIAQVANSY-PMIWMTEARDRAGNITVERGYLMEEGPNSFOSDP 132
DB 62 VGGISIGICIAQALATKHPDAAPLIVTEAKDRGGINITREENGFLMEGPNFSFOSDP 121
QY 133 MLTMAVDCGLKDDLVLDGPNAPRVVLVKGKLRPVPSKLTDLFFDLMSIPKLRAGFPI 192
DB 122 MLTMAVDSGLKDDLVLDGPNAPRVVLVKGKLRPVPSKLTDLFFDLMSIPKLRAGFPI 181
QY 193 GLRSPPGHESVDFPVRNIGGVFRLTEPPSGYVVDPSKLSKAAFGKYWKLEET 252
DB 182 GIRSPPGHESVDFPVRNIGGVFRLTEPPSGYVVDPSKLSKAAFGKYWKLEON 241
QY 253 GSIIGGTFAIKERSSTPKAPRDPRLPKPGQTVGSFRKGLRLPAISARLSKSLKS 312
DB 242 GSIIGGTFAIKERSSTPKAPRDPRLPKPGQTVGSFRKGLRLPAISARLSKSLKS 301

QY 313 WKLSITSEKGYHLTYETPEGVVSIOSRSIWTVPYSVANSILRPLSYAADALSNEY 372
DB 302 WKLSITSEKGYHLTYETPEGVVSIOSRSIWTVPYSVANSILRPLSYAADALSNEY 361
QY 373 YPPVAVTISTPOBAIDERLVDELKFGQLHPRQGVETLTIVSSSLFPNRAKGRV 432
DB 362 YPPVAVTISTPOBAIDERLVDELKFGQLHPRQGVETLTIVSSSLFPNRAKGRV 421
QY 433 LLNVYIGAKNPETLSTESQIVYVNDRLKMLIKPKADPLVGVRYVPAIPQFLVG 492
DB 422 LLNVYIGAKNPETLSTESQIVYVNDRLKMLIKPKADPLVGVRYVPAIPQFLVG 481
QY 493 HLDLTSTAKAMNDNGLEGLFLGNYVSVAGALRCVAGAYEVSAYGFLSRYAK 548
DB 482 HLDLTSTAKAMNDNGLEGLFLGNYVSVAGALRCVAGAYEVSAYGFLSRYAK 537

RESULT 7
US-08-808-323-2
Sequence 2, Application US/0808323
Patent No. 6018105
GENERAL INFORMATION:
APPLICANT: Johnson, Marie
APPLICANT: Volrath, Sandra
TITLE OF INVENTION: Promoters from Plant
TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: No. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,323
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-808-323-2
Query Match 73.0%; Score 2075; DB 3; Length 537;
Best Local Similarity 75.4%; Pred. No. 5.7e-198;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

1 TITLE OF INVENTION: OXIDASE ("PROTOX")
2 NUMBER OF SEQUENCES: 43
3 CORRESPONDENCE ADDRESSES:
4 ADDRESSEE: No. 6084155artis Corporation
5 STREET: 3054 Cornwallis Road
6 CITY: Research Triangle Park
7 STATE: NC
8 COUNTRY: USA
9 ZIP: 27709
10
11 COMPUTER READABLE FORM:
12 MEDIUM TYPE: Floppy disk
13 COMPUTER: IBM PC compatible
14 OPERATING SYSTEM: PC-DOS/MS-DOS
15 SOFTWARE: Patentin Release #1.0, Version #1.30
16
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/09/102,420B
19 FILING DATE: 22-JUN-1998
20 CLASSIFICATION: 800
21
22 PRIOR APPLICATION DATA:
23 APPLICATION NUMBER: US 09/059,164
24 FILING DATE: 13-APR-1998
25
26 PRIOR APPLICATION DATA:
27 APPLICATION NUMBER: US 09/050,603
28 FILING DATE: 30-MAR-1998
29
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US 60/126,430
32 FILING DATE: 11-MAR-1998
33
34 PRIOR APPLICATION DATA:
35 APPLICATION NUMBER: US 08/808,931
36 FILING DATE: 28-FEB-1997
37
38 PRIOR APPLICATION DATA:
39 APPLICATION NUMBER: US 60/012,705
40 FILING DATE: 28-FEB-1996
41
42 PRIOR APPLICATION DATA:
43 APPLICATION NUMBER: US 60/013,612
44 FILING DATE: 28-FEB-1996
45
46 PRIOR APPLICATION DATA:
47 APPLICATION NUMBER: US 60/020,003
48 FILING DATE: 21-JUN-1996
49
50 PRIOR APPLICATION DATA:
51 APPLICATION NUMBER: US 08/472,028
52 FILING DATE: 06-JUN-1995
53
54 ATTORNEY/AGENT INFORMATION:
55 NAME: Meigs, J. Timothy
56 REGISTRATION NUMBER: 38,241
57 REFERENCE/DOCKET INFORMATION:
58 TELEPHONE: (919) 541-8587
59 TELEFAX: (919) 541-8587
60 INFORMATION FOR SEQ ID NO: 2:
61 SEQUENCE CHARACTERISTICS:
62 LENGTH: 537 amino acids
63 TYPE: amino acid
64 TOPOLOGY: linear
65 MOLECULE TYPE: protein
66 US-09-102-420B-2

Query Match 73.0%; Score 2075; DB 3; Length 537;
Best Local Similarity 75.4%; Pred. No. 5.7e-198;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

QY 22 LAFLNRTSFPISRSKSNVNCNM--RTRCSYAKDYTPSSAVDGGAAEL--PCVI 76
DB 3 LSLRRTTQSLPFSKPR-LRLANYKRLRLRCVAGSPYVGSSTIRGGGTTITTCVI 61
QY 77 VAGISGLCIAVMSANT---PNIMATVENDRAGNITTVREGRYLMEEGPNSEFSDP 132
DB 62 VEGGISGLICIAALATKHPDAPNLIVTEAKDRVGNITTEENGFLMBEGPNSEFSDP 121
QY 133 MLTMAVDCGLKDDLVLGPNAPRFVLMKGLRPVPSKITDLPFDLMSIRCKLRAGGPI 192
DB 122 MLTMAVDCGLKDDLVLGPTAPRFVLMKGLRPVPSKITDLPFDLMSIRCKLRAGGPI 181

QY 193 GLRSPPECHESVQFYRNRLGVEFERLLIEPFCGYYVGDPSKLSMKAPGVMKLEET 252
DB 182 GIRSPPECHESVEEYFRNRINGDEVEFERLLIEPFCGYYADPDKLSMKAPGVMKLEON 241
QY 253 GGSIIIGTFEKAIRKERSSTPKAPRDPRLPKPKGOTVGSFRGLRMLDPAISARLSGLKLS 312
DB 242 GGSIIIGTFEKAIRKAPRDPRLPKPKGOTVGSFRGLRMLDPAISARLSGLKLS 301
QY 313 WKLSITKSEKGYHLYTTEPEGVSLQSRSIYMTVPSYASNLRLPSVAADALSNEY 372
DB 302 WKLSITKSEKGYHLYTTEPEGVSLQSRSIYMTVPSYASNLRLPSVAADALSNEY 361
QY 373 YPPGATITVPOFATIDELVNGELKRGQALPRPQGVETLGTIYSSSLFPNRPKGRV 432
DB 362 YPPGATITVPOFATIDELVNGELKRGQALPRPQGVETLGTIYSSSLFPNRPKGRV 421
QY 433 LLNVIYGANPRLKTESQLEVEVDRLRKLIFPKADDPVGVGVWPAIFQPLVG 492
DB 422 LLNVIYGANPRLKTESQLEVEVDRLRKLIFPKADDPVGVGVWPAIFQPLVG 481
QY 493 HLDLSTAKAANDNGLEGLFCGNVYSGVALGRCEGAYEVAEVTGLSRVAYK 548
DB 482 HFDLDTAKSSLTSSGVEGLFGLGNVYAGVALGRCEGAYETAIEVNNMRSRYAYK 537

RESULT 10
US-09-071-296-2
Sequence 2, Application US/09071296
Patent No. 6177245
GENERAL INFORMATION:
APPLICANT: Ward, Eric R
TITLE OF INVENTION: Manipulation of Protoporphyrinogen
TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,296
FILING DATE: 06-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/261,198
FILING DATE: 16-JUN-94
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 537 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-296-2

Query Match 73.0%; Score 2075; DB 4; Length 537;
Best Local Similarity 75.4%; Pred. No. 5.7e-198;

[illegible]

```

1      APPLICATION NUMBER: US 60/013,612
2      FILING DATE: 28-FEB-1996
3      PRIOR APPLICATION DATA:
4      APPLICATION NUMBER: US 60/020,003
5      FILING DATE: 21-JUN-1996
6      ATTORNEY/AGENT INFORMATION:
7      NAME: Meigs, J. Timothy
8      REGISTRATION NUMBER: 38,241
9      REFERENCE/DOCKET NUMBER: CGC 1847
10     TELECOMMUNICATION INFORMATION:
11     TELEPHONE: (919) 541-8587
12     TELEFAX: (919) 541-6889
13     INFORMATION FOR SEQ ID NO: 12:
14     SEQUENCE CHARACTERISTICS:
15     LENGTH: 543 amino acids
16     TYPE: amino acid
17     TOPOLOGY: linear
18     MOLECULE TYPE: protein
19     US-08-008-931-12

```

| | | | | |
|-----------------------|---------|---------------------|--------|----------------|
| Query Match | 72.38; | Score 2055; | DB 2; | Length 543; |
| Best Local Similarity | -78.18; | Pred. No. 5.7e-196; | | |
| Matches | 392; | Conservative | 44; | Mismatches 60; |
| | | | Indels | 6; |
| | | | Gaps | 3; |

[illegible]

| | | | |
|----|-----|--|-----|
| Qy | 107 | GCNITVERDXYLWEEGNSNPQSDPMLTMAVOCGLKDDLVLPBPNAFRCVLWKGKLRPV | 166 |
| Dp | 102 | GCNITTHMRDGYLWEEGNSNPQSDPMLTMYVDSGLKDELVLGDPDPAFRLVLMNRKLRPV | 161 |
| Qy | 167 | PSKLTLDJFFDLMSIPGKLRAGFCRPIGLRSPSPGHEESBOFVARNLGCSEVERLIEPFC | 226 |
| Dp | 162 | PGKLTLDJFFDLMSIGCKIRAGFCALIRPPPHHEESVEEFVARNLGDSEVERLIEPFC | 221 |
| Qy | 227 | SGVYVDPDSKLSMAAFCGVKWKLEETGGSIIGCFEFAIKERSSRPKARPRDRLPKPGQT | 286 |
| Dp | 222 | SGVYAGDSPSKLSMAAFCGVKWKLEKNGSIIIGCFEFAIQENNGASKPRDRLPKPGQT | 281 |
| Qy | 287 | VGSFPRKGLRMLPDAISARLGSKLSLWKLSITKSEKAGHLYETPEGVVSLQSRITVM | 346 |
| Dp | 282 | VGSFPRKGLTMLPDAISARLGNKVKLTKMSKLSISKLDGSEVSLYETPEGVVSLQCKITVVL | 341 |
| Qy | 347 | TVPSYVASNILRPLSVAAADALSNFYTPPGAVTISTPQEAIRDERLYVDELFKFGQLHP | 406 |
| Dp | 342 | TIPSYVASTLRLRPISAAADALSKFYPPVAAVSI SYPKERIRRECDLIDELKFGQLHP | 401 |
| Qy | 407 | RTQGEYETIGTYSLSLFPNRPAPKRVLLANTYGAKKPELISKTESOLIVEVBRDLEKML | 466 |
| Dp | 402 | RSQGEYETIGTYSLSLFPNRPAPRVLLANTYGATWTGILSKDSLSLVEYVBRDLKIL | 461 |
| Qy | 467 | IKPKAODPLVGVVRVWPOLIQFLVGHLDLTSTAKAAMNDNGLEGLFLGGNYYSGVALGR | 526 |
| Dp | 462 | INPNADOPFVVGVRILMPOLIQFLVGHLDLDVAKASIRNTGEGEGLFLGGNYYSGVALGR | 521 |
| Qy | 527 | CVEGAYEVASEVTGFLSNRYKX | 548 |
| Dp | 522 | CVEGAYEVAAEYNDFLTNRYKX | 543 |

RESULT 14
US-09-102-420B-12
Sequence 12 Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volrath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
NUMBER OF INVENTION: OXIDASE ("PROFOX")
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931

```

1      FILING DATE: 28-FEB-1997
2      PRIOR APPLICATION DATA:
3      APPLICATION NUMBER: US 60/012,705
4      FILING DATE: 28-FEB-1996
5      PRIOR APPLICATION DATA:
6      APPLICATION NUMBER: US 60/013,612
7      FILING DATE: 28-FEB-1996
8      PRIOR APPLICATION DATA:
9      APPLICATION NUMBER: US 60/020,003
10     FILING DATE: 21-JUN-1996
11     PRIOR APPLICATION DATA:
12     APPLICATION NUMBER: US 08/472,028
13     FILING DATE: 06-JUN-1995
14     ATTORNEY/AGENT INFORMATION:
15     NAME: Meigs, J. Timothy
16     REGISTRATION NUMBER: 38,241
17     REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
18     TELECOMMUNICATION INFORMATION:
19     TELEPHONE: (919) 541-8587
20     TELEFAX: (919) 541-8689
21     INFORMATION FOR SEQ. ID NO.: 12:
22     SEQUENCE CHARACTERISTICS:
23     LENGTH: 543 amino acids
24     TYPE: amino acid
25     TOPOLOGY: linear
26     MOLECULE TYPE: protein
27     US-09-102-420B-12

```

| | | | | |
|-----------------------|------------------|---------------------|-----------|-------------|
| Query Match | 72.3%; | Score 2055; | DB 3; | Length 543; |
| Best Local Similarity | 78.1%; | Pred. No. 5.7e-196; | | |
| Matches 392; | Conservative 44; | Mismatches 60; | Indels 6; | Gaps 3; |

[illegible]

RESULT 15
US-08-808-931-20
: Sequence 20, Application US/08808931

Patent No. 5939602
GENERAL INFORMATION:
APPLICANT: Volirath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Potter, Sharon
APPLICANT: Ward, Eric
APPLICANT: Heiletz, Peter
TITLE OF INVENTION: DNA Molecules Encoding Plant
Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5939602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 536 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-808-931-20

Query Match 70.6%; Score 2006.5; DB 2; Length 536;
Best Local Similarity 73.1%; Pred. No. 3.8e-191;
Matches 391; Conservative 52; Mismatches 83; Indels 9; Gaps 4;

QY 22 LAFINTSTI-PESSISKRNVCNGWTRCSYAKDYTPSSAVDGGPAE--LDCVIV 77
DB 3 LSLRQPLSPFSPNPPRR-SRPYKPLMLKCSVSGSVGSSTIEGGGGKTVTADCVIV 61
QY 78 GAGISGLCIAQYMSANYP---NLMTYARDRAGNITTYERDGYLMEEGPNSFQSPDPM 133
DB 62 GGISGLCIAQALVTKRHDPAKNVYVTEAKDRVGNITITREOGFLMEEGSPNSFQSPDPM 121
QY 134 LTMAYVDCGLKDLVGDNDNAPRFVLMKGLRPVPSKLTDLPPFDLMSIPGKLRAGFGPTG 193
DB 122 LTMVYDSGLKDLVGLDPTAPRFVLMKGLRPVPSKLTDLPPFDLMSIGKIRAGFGAIG 181
QY 194 LRSPPGHEESYEQFVRNRLGGEVFERLIEPFCGSGVYVGDPSKLSMKAAFGKVMKLEETG 253
DB 182 IRSPPGRESEYEEFVRNRLGDEVFERLIEPFCGSGVYVADPAKLSMKAAFGKVMKLEENG 241

QY 254 GSIIIGTFPAIKERSSTPKAPRDPRLPKKGQVGSFRRGLRMLPDATISARLSKLTLSW 313
DB 242 GSIIIGAFPAIQAANKKAPRTDPRLPKRGQVGSFRRGLRMLPEATISARLGDGVSW 301
QY 314 KLSSTIKSEKSGYHCLTYETPEBEGVSLQGRSIVMYPSVYANILIRPSVANAADLSNFY 373
DB 302 KLSSTIKLASGESLYETPEBEGVIVQSKSVMMYPSHVASSLIRPLSDNAEKLKLY 361
QY 374 PPGAVTISYPOFAIRDERLVDELKGFGLHPTQGVETLGTIYSSSLFPNNAPKGRVL 433
DB 362 PVAASISYAKAIRSECLIDGELKGFGLHPTQKVELTGTIYSSSLFPNNAPKGRVL 421
QY 434 LMTYGAKKPELSTESQVLEVVDRLRKMLIKPKAODPLVGVYVWQATPQFLVGH 493
DB 422 LMTYGATNTGILSKSEBELVEAVDRDLRKMLIKPSTDPVLVGVKLMQALPQFLIGH 481
QY 494 LDTLSTAKAAMNDNGLEGLFLGNYVSGVALGRCEGAYEVAASEVTGFLSRVAYK 548
DB 482 IDLVDAKAKSLSSSGHGLFLGNYVAGVALGRCEGAYETATQVNDPMGRVAYK 536

Search completed: July 3, 2001, 10:31:54
Job time: 1913 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: July 3, 2001, 10:33:07 ; Search time 34.09 Seconds

(without alignments)
550.660 Million cell updates/sec

Title: US-09-508-418-2

Sequence score: 2844
1 MTTPTIANNHPIFTHOSSSS.....EGAEVASEVTGELSRVAVK 548

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | Match Length | ID | Description |
|------------|-------|-------|--------------|----|-------------|
| 1 | 2821 | 99.2 | 548 | 1 | PROOC_TORAC |
| 2 | 2075 | 73.0 | 537 | 1 | PROOC_ARATH |
| 3 | 541 | 19.0 | 471 | 1 | PROOX_MYXA |
| 4 | 505 | 17.8 | 504 | 1 | PROOX_TORAC |
| 5 | 485 | 17.1 | 470 | 1 | PROOX_BACSU |
| 6 | 375.5 | 13.2 | 527 | 1 | PROOX_PROFR |
| 7 | 363.5 | 12.8 | 477 | 1 | PROOX_HUMAN |
| 8 | 362.5 | 12.7 | 477 | 1 | PROOX_MOUSE |
| 9 | 302.5 | 10.6 | 490 | 1 | YAM7_SCHPO |
| 10 | 221 | 7.8 | 426 | 1 | PROOX_MYCTU |
| 11 | 216.5 | 7.6 | 451 | 1 | PROOX_MYCLE |
| 12 | 189 | 6.6 | 539 | 1 | PROOX_YEAST |
| 13 | 141 | 5.0 | 448 | 1 | AOFR_MYCTU |
| 14 | 139 | 4.9 | 570 | 1 | CRTI_MARPS |
| 15 | 134 | 4.7 | 522 | 1 | AOFA_ONCMY |
| 16 | 132.5 | 4.7 | 527 | 1 | AOFA_BOVIN |
| 17 | 131 | 4.6 | 566 | 1 | CRTI_ARATH |
| 18 | 125.5 | 4.4 | 417 | 1 | YVAB_RHISN |
| 19 | 125.5 | 4.4 | 472 | 1 | CRTI_SYNY3 |
| 20 | 125.5 | 4.4 | 489 | 1 | CBPL_CANAL |
| 21 | 123.5 | 4.3 | 526 | 1 | AOFA_RAT |
| 22 | 123 | 4.3 | 582 | 1 | CRTI_CAPAN |
| 23 | 122 | 4.3 | 471 | 1 | PVO_MICRU |
| 24 | 119.5 | 4.2 | 478 | 1 | PVO_SYNY3 |
| 25 | 118 | 4.1 | 571 | 1 | CRTI_MAIZE |
| 26 | 117.5 | 4.1 | 469 | 1 | P49_STRLI |
| 27 | 116 | 4.1 | 583 | 1 | CRTI_LYCES |
| 28 | 114.5 | 4.0 | 1106 | 1 | ITR7_RAT |
| 29 | 114 | 4.0 | 531 | 1 | ACHC_ACHRU |
| 30 | 112.5 | 4.0 | 570 | 1 | CRTI_SOYBN |
| 31 | 111 | 3.9 | 508 | 1 | CRTI_STRSE |
| 32 | 109.5 | 3.9 | 527 | 1 | AOFA_HUMAN |
| 33 | 109.5 | 3.9 | 908 | 1 | SRCA_RABIT |

| | | | | | | |
|----|-------|-----|------|---|------------|---------------------|
| 34 | 109 | 3.8 | 1181 | 1 | ITR7_HUMAN | Q13663 homo sapien |
| 35 | 108.5 | 3.8 | 1062 | 1 | CERU_MOUSE | Q61147 mus musculus |
| 36 | 108 | 3.8 | 504 | 1 | A37C_DROME | P18487 drosophila |
| 37 | 107.5 | 3.8 | 516 | 1 | ER13_ARATH | Q65403 arabidopsis |
| 38 | 107 | 3.8 | 405 | 1 | SAOX_CORSI | P40875 corynebacte |
| 39 | 107 | 3.8 | 1179 | 1 | ITR7_MOUSE | Q61738 mus musculus |
| 40 | 106.5 | 3.7 | 630 | 1 | ITR1_MOUSE | O09046 mus musculus |
| 41 | 105.5 | 3.7 | 872 | 1 | IP3L_RAT | P42335 rattus norv |
| 42 | 105 | 3.7 | 755 | 1 | TR2M_AGR3 | P25017 agrobacteri |
| 43 | 104.5 | 3.7 | 466 | 1 | AMID_PSEPU | O69768 pseudomonas |
| 44 | 104.5 | 3.7 | 517 | 1 | ER12_ARATH | Q65402 arabidopsis |
| 45 | 104.5 | 3.7 | 524 | 1 | CRT1_RHOCA | P17054 rhodobacter |

ALIGNMENTS

| RESULT | 1 | STANDARD: | PRT: | 548 AA. |
|-------------|--|-----------|-----------|-----------------------------|
| PROOC_TORAC | | | | |
| AC | 024163: | | | |
| DT | 15-DEC-1998 (Rel. 37, Created) | | | |
| DT | 15-DEC-1998 (Rel. 37, Last sequence update) | | | |
| DT | 15-JUL-1999 (Rel. 38, Last annotation update) | | | |
| DE | PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR (EC 1.3.3.4) (PPO I) | | | |
| DE | (PROTOPORPHYRINOGEN IX OXIDASE ISOZYME I) (PPX I). | | | |
| GN | PPXI. | | | |
| OS | Nicotiana tabacum (Common tobacco). | | | |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; | | | |
| OC | Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; | | | |
| OC | Solanales; Solanaceae; Nicotiana. | | | |
| OX | NCBI_TaxID=4097; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=CV. SRI. | | | |
| RX | MELINE=97385200; PubMed=9238074; | | | |
| RA | Lemontova I., Kruse E., Mock H.-P., Grimm B.; | | | |
| RT | "Cloning and characterization of a plastidial and a mitochondrial | | | |
| RT | isoform of tobacco protoporphyrinogen IX oxidase."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997). | | | |
| CC | -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN | | | |
| CC | IX TO FORM PROTOPORPHYRIN IX. | | | |
| CC | -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN- | | | |
| CC | IX + H(2)O(2). | | | |
| CC | -1- COFACTOR: FAD (BY SIMILARITY). | | | |
| CC | -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS, AND | | | |
| CC | CHLOROPHYLL SYNTHESIS. | | | |
| CC | -1- SUBUNIT: HOMODIMER; CONTAINS ONE FAD PER HOMODIMER (BY | | | |
| CC | SIMILARITY). | | | |
| CC | -1- SUBCELLULAR LOCATION: CHLOROPLAST. | | | |
| CC | -1- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES. | | | |
| CC | DECREASED EXPRESSION IN OLDEST LEAVES. EXPRESSED AT VERY LOW LEVEL | | | |
| CC | IN ROOTS. | | | |
| CC | -1- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL | | | |
| CC | EXPRESSION IN THE DARK PERIOD. | | | |
| CC | -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyrighted. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@sib-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL: Y13465; CAA73865.1; - | | | |
| KW | Protoporphyrin biosynthesis; Heme biosynthesis; Oxidoreductase; | | | |
| KW | Transit peptide; Flavoprotein; FAD; Chloroplast; | | | |
| KW | Chlorophyll biosynthesis. | | | |
| FT | TRANSIT | 1 | 50 | CHLOROPLAST (POTENTIAL). |
| FT | CHAIN | 51 | 548 | PROTOPORPHYRINOGEN OXIDASE. |
| FT | NP_BIND | 78 | 83 | FAD (POTENTIAL). |
| FT | SEQUENCE | 548 AA; | 59230 MW; | 66892E78FB8A3EE30 CRC64; |

Query Match 73.0%; Score 2075; DB 1; Length 537;
 Best Local Similarity 75.4%; Pred. No. 3,1e-149;
 Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

22 LAFILRTSEFIPSSISKSNVNCNM---KRCISAKIYTPSSAVDGGPAEL-DCVI 76
 3 USLMPRTOSLSPFSKSN-LKLNRYKPLRLRCVAGGPVSSKTEGGGTTTITTCVI 61

77 VGAGISGLCIAQVMSANP-----PMLVTEARDRAGANTTTERDGYLMEEGPNSFOPSD 132
 62 VGGISGLICIAQALATKHPDAPNLIVTEADNRGNITTEENGFLMEEPNSFQSDP 121

133 MLTMAVDGKADLVLGPPNAPRYLWKAKIRPVPSKITDLPFDLMSIPKLRAGCPI 192
 122 MLTMAVDGKADLVLGPPNAPRYLWKAKIRPVPSKITDLPFDLMSIPKLRAGCPI 181

193 GLRPSPEHSESVDFVRNLTNGEVEERLTIEPFGSGVYVGGPSKLSMKAAGKWKLEET 252
 182 GTRSPSPRESSEVEFRVARNLGDVEFEERLTIEPFGSGVYVGGPSKLSMKAAGKWKLEON 241

253 GSGIIGTFKAIKERSSTPKAPRDLPEKQTVSGFRKGLRLMLPDAISARLSKTLKS 312
 242 GSGIIGTFKAIKERSSTPKAPRDLPEKQTVSGFRKGLRLMLPDAISARLSKTLKS 301

313 WKLSITSEKGYHLIYETPEGVYSLQSNSTVMTVPSYVANSILRLPSVAADALSNEY 372
 302 WKLSITSEKGYHLIYETPEGVYSLQSNSTVMTVPSYVANSILRLPSVAADALSNEY 361

373 YPVGAVTISYPOAIRDERLVDELKGFQGLHPTQGVETLITVSSSLFPNAPRGV 432
 362 YPVGAVTISYPOAIRDERLVDELKGFQGLHPTQGVETLITVSSSLFPNAPRGV 421

433 LLNTIGGAKNPELISKTESQVLEVVNDLRKMLIKPKADPLVGVVYVPAIQPLVG 492
 422 LLNTIGGAKNPELISKTESQVLEVVNDLRKMLIKPKADPLVGVVYVPAIQPLVG 481

493 HLDITSTAKAMNDNGLEGLFGNIVYSGVALGRCVGAATEVASEVGEISRAVYK 548
 482 HFDITSTAKAMNDNGLEGLFGNIVYSGVALGRCVGAATEVASEVGEISRAVYK 537

RESULT 3
 ID PROX MYXXA STANDARD; PRT; 471 AA.
 AC P56601;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN HEMV.
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
 OC Myxococcales; Cystobacteriineae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=92192453; PubMed=1372277;
 RA Li S., Lee B.U., Shimkets L.J.;
 RT "csga expression entrains Myxococcus xanthus development.";
 RL Genes Dev. 6:401-410(1992).
 RN [2]
 RP REVISION, AND CHARACTERIZATION.
 RP MEDLINE=96224079; PubMed=8621504;
 RA Dalley H.A., Dalley T.A.;
 RT "Protoporphyrinogen oxidase of Myxococcus xanthus. Expression,
 purification, and characterization of the cloned enzyme.";
 RL J. Biol. Chem. 271:8714-8718(1996).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 IX TO FORM PROTOPORPHYRIN IX. MAY HAVE AN ESSENTIAL EXTRACELLULAR
 FUNCTION DURING DEVELOPMENT POSSIBLY AS A PHEROMONE.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 IX + H(2)O(2).

CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND POMPORPHIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
 CC FRAMESHIFT AT POSITION 127.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, W3709; NOT_ANNOTATED_CDS.
 CC DR EMBL; AF098938; AAD13609.1; -
 CC For phylin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
 CC Heme biosynthesis.
 CC FT NE_BIND 16 21 FAD (PROTENTIAL)
 CC SO SEQUENCE 471 AA; 49387 MW; 2E9909D464F81515 CRC64;

Query Match 19.0%; Score 541; DB 1; Length 471;
 Best Local Similarity 31.7%; Pred. No. 1.9e-33;
 Matches 150; Conservative 84; Mismatches 207; Indels 32; Gaps 12;

72 LDCVYAGISGLCIAQVMSANPMLVTEARDRAGANTTTERDGYLMEEGPNSFOPSD 131
 10 MNVAVVGSGISGLAVAHNHRSGTDAVLLESARLGAAGTALAGYVEGQPNFSLRE 69

132 PML-TMAVDCGKDDLVLDGPNAP-REVLKMGTRPVPSKITDLPFDLMSIPKLRAGC 189
 70 PATRALAALNLEGRIRAADPAKRRVYTRGRLSVSPAPFASDILPLGARLRVA- 128

190 GPGILRSPPEHSESVDFVRNLTNGEVEERLTIEPFGSGVYVGGPSKLSMKAAGKWKLE 249
 129 GELFSRRAEAGEVDSLAAGRHLGHRAQVLLDAVQGTIVAGVDEQSLVATFPMVKM 188

250 EETGSGIIGTFKAIKERSSTPKAPRDLPEKQTVSGFRKGLRLMLPDAISARLS 305
 189 EREHNSLILGATNA-----OKAQRQALPAGTAPKLSGALSTGDDGQVLDLALASL 241

306 GSKLTKLMSITSEKGYHLIYETPEGVYSLQSNSTVMTVPSYVANSILRLPSVAAD 365
 242 GDAHVAGARVEGLAR-EDSGMRLIIEHGRRAELSYAGVYLAAPAHATKLRPLDDLA 300

366 DALSNFYTPVGAVTISYPOAIRDERLVDEL---KGFQGLHPTQGVETLITVSSSL 422
 301 ALVAGIAYAPVIAVHVLGFDA-----GTLPAPDGGLVPAEQRRLMGAITHASTT 350

423 FPNRAPKGVLLNTYGGKKNPELISKTESQVLEVVNDLRKMLIKPKADPLVGVVYV 482
 351 FPNRAPKGVLLNTYGGKKNPELISKTESQVLEVVNDLRKMLIKPKADPLVGVVYV 408

483 POALPOFLVGHLDITSTAKAMNDNGLEGLFGNIVYSGVALGRCVGAATEVASEV 535
 409 PLGIPOYNLGHLEVAALDAL--QRLPGHLIGNAYGVGLNDCIRNAADLA 459

RESULT 4
 ID PROX TOBAC STANDARD; PRT; 504 AA.
 AC O24164;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE, MITOCHONDRIAL (EC 1.3.3.4) (PPO II)
 GN PPXII OR PPX2.
 OS Nicotiana glauca (Common tobacco).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta: eudicotyledons: core eudicots: Asteridae: euasterids I:
 OC Solanales: Solanaceae: Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SRL;
 RX MEDLINE=97385200; PubMed=9238074;
 RA Lermontova I., Kruse E., Mock H.-P., Grimm B.;
 RT "Cloning and characterization of a plastidal and a mitochondrial
 isoform of tobacco protoporphyrinogen ix oxidase."
 RL Proc. Natl. Acad. Sci. U.S.A. 94:8895-8900(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. Samsun NN;
 RX PubMed=9808719;
 RA Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
 RA Isogai A.;
 RT "Molecular characterization of photomixotrophic cultured tobacco cells
 resistant to protoporphyrinogen oxidase-inhibiting herbicides."
 RL Plant Physiol. 118:751-758(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SRL;
 RA Horikoshi M., Mametsuka K., Hirooka T.;
 RT "The molecular basis of photobleaching herbicide resistance in
 tobacco."
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 IX TO FORM PROTOPORPHYRIN IX.
 CC -1- FUNCTION: PROVIDES PRECURSOR FOR THE MITOCHONDRIAL AND PLASTIDIC
 HEME SYNTHESIS AND THE PREDOMINANT CHLOROPHYLL SYNTHESIS IN
 PLASTIDS.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) =
 PROTOPORPHYRIN-IX + H(2)O(2).
 CC -1- COFACTOR: FAD (BY SIMILARITY).
 CC -1- ENZYME REGULATION: INHIBITED BY THE HERBICIDE ACIFLUOREN.
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRION.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EXPANDING PREMATURE LEAVES.
 CC -1- DECREASED EXPRESSION IN OLDEST LEAVES. ALSO DETECTED IN ROOTS.
 CC -1- INDUCTION: OSCILLATING EXPRESSION DURING DIURNAL GROWTH. MAXIMAL
 EXPRESSION IN THE DARK PERIOD.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y13466; CAA73866.1;
 DR EMBL: AB020500; BAA34712.1;
 DR EMBL: AF044129; AAD02291.1;
 DR InterPro: IPR002937;
 DR Pfam: PF01593; Amino oxidase; 1.
 DR Prophylin biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Mitochondrion.
 FT NP_BIND 20 25 FAD (POTENTIAL).
 SQ SEQUENCE 504 AA; 55407 MW; B85B55EC81DC00A CRC64;

Db 62 WDEGANTMTSESGDVTFILDSIGLREKQFPLSQNKRRIARNGPVLPLSPNIDLIKSNE 121
 QY 179 MSIFGKLKAGCGPI-----GLRSPFGHESYEDFVARNRNGGEFFELIIPFGSGYVGD 233
 Db 122 LSTGSKLOMLEPILMKRNKLSQVSDSH-ESVSGFQRHFGKEVVDLIDPEVAGTCGGD 180
 QY 234 PSKLSMKAAPGKWKLEPTGSIIGTFEKA-IKERSSTPAAPRDPRLPKPGQTVGSFRK 292
 Db 181 PDSLSMHHSFPELWMLNKRFSVILGALRSLRNKKNKKQCPRTSANKRKQSGSFFLG 240
 QY 293 GLRMLPDAI-----SARLGSK-LKL-----SMKLSITSEKGGHILYETP 333
 Db 241 GMQTLTDAICDRLREDELRLNSRVLELSCSTEDSAIDSSIIASPHKQSEESFD- 298
 QY 334 EGVVSLGSRISVIVPVSVASNIILRPLSVAAADLSNFY-----YPPGATISYPOPAIR 389
 Db 299 -----AVIMTAPLIDVKS- -KIARRGNPLNLTPEVDVPLSVITTTKRENVK 347
 QY 390 DERLVDGLKGFOLHP---RTQGVETLGTYSLSLFPNRAKGRVLLNLYGAKNPEI 446
 Db 348 -----YLEGFGVLYPSKEQHGKLTGLTFSSMPPDRAPNNVLYTFVGSGRRREL 401
 QY 447 LKTESQLEVEYVDRDLRKMLIKPKAQDPLVGVGVWPAIPQVLGH-LDTLSTAKAAN 505
 Db 402 AKASRTELKEIYTSDLKQLL--GAEGEPTYVNHLYWSKAFP--LYGHNYDSVLDAIDKME 457
 QY 506 DNGLEGLFLGNGVSGVALGRCEAGYEVASEVTGFL 542
 Db 458 KN-LPGLFYAGNHNGLSVGNALSSGCNADLVISYL 493

RESULT 5
 PROX_BACSU STANDARD: PRT: 470 AA.
 ID PROX_BACSU
 AC P32397;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PRO).
 DE HEWY OR HEMG.
 GN Bacillus subtilis.
 OS Bacillus; Firmicutes; Bacillus/Clostridium group;
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93094140; PubMed=1459957;
 RA Hanson M., Hederstedt L.;
 RT "Cloning and characterization of the Bacillus subtilis hewy gene
 cluster, which encodes protoheme IX biosynthetic enzymes."
 RL J. Bacteriol. 174:8081-8093(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;
 RT Submitted (JUN-1997) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=94117488; PubMed=828631;
 RA Dailey T.A., Meissner P., Dailey H.A.;
 RT "Expression of a cloned protoporphyrinogen oxidase."
 RL J. Biol. Chem. 269:813-815(1994).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 IX TO FORM PROTOPORPHYRIN IX. ALSO OXIDIZES THE PATHWAY
 INTERMEDIATE COPROPORPHYRINOGEN III.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 IX + H(2)O(2).
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
 CC -1- PATHWAY: PENULTIMATE STEP IN PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----

Query Match 17.8%; Score 505; DB 1; Length 504;
 Best Local Similarity 28.8%; Pred. No. 1.1e-30;
 Matches 149; Conservative 100; Mismatches 208; Indels 60; Gaps 17;

OY 60 PSSAVDGGPAELDCVYAGISGICIAOVMSANYPNLMVTEARDRAGNITTYERDGYL 119
 Db 3 PSAGEDKHSKAK-NVAVIGAGVSGLAAYAKRKIKGLVWTFEAGKAGKRLRSVSGGLI 61
 OY 120 WEEGPNFSQPSDPMLTMAVD-CGLKDLVLGCDPNAPRFVLMKGLRVPKSLDLPFDL 178

| | | | | |
|---|--|---|-----------------------------|--------------|
| OY | 494 | LDTLSTAKAAMDNGL---- | EGLF LGAGVSCALGRCEGAEVAASEY | 538 |
| | : | ::::: | : : : : | |
| D6 | 460 | LE-----RAAVVDSTLAQRPTMAVAGSALHGLPDCISDRHSADSV | 503 | |
| RESULT 7 | | | | |
| PPOX_HUMAN | ID | PROX_HUMAN | STANDARD: | PRT: 477 AA. |
| AC | P50336; | | | |
| DT | 01-OCT-1996 | (Rel. 34, Created) | | |
| DT | 01-OCT-1996 | (Rel. 34, Last sequence update) | | |
| DT | 15-DEC-1998 | (Rel. 37, Last annotation update) | | |
| DE | PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO). | | | |
| GN | PPOX. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBII_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Placenta; | | | |
| RX | MEDLINE=95229621; PubMed=7713909; | | | |
| RA | Nishimura K., Taketani S., Inokuchi H. ; | | | |
| RT | "Cloning of a human cDNA for protoporphyrinogen oxidase by | | | |
| RT | complementation in vivo of a hemg mutant of Escherichia coli." | | | |
| RL | J. Biol. Chem. 270:8076-8080(1995). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Placenta; | | | |
| RX | MEDLINE=96367087; PubMed=87711201; | | | |
| RA | Dalley T.A., Dalley H.A.; | | | |
| RT | "Human protoporphyrinogen oxidase: expression, purification, and | | | |
| RT | characterization of the cloned enzyme."; | | | |
| RL | Protein Sci. 5:98-105(1996). | | | |
| RN | [3] | | | |
| RP | VARIANT VP ARG-232, AND VARIANT HTS-304. | | | |
| RX | MEDLINE=97005368; PubMed=8852667; | | | |
| RA | Deebach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V., | | | |
| RA | Grandchamp B., Nordmann Y.; | | | |
| RT | "Mutations in the protoporphyrinogen oxidase gene in patients with | | | |
| RT | variegate porphyria."; | | | |
| RL | Hum. Mol. Genet. 5:407-410(1996). | | | |
| RN | [4] | | | |
| RP | VARIANT VP TRP-59, AND VARIANT CYS-168. | | | |
| RX | MEDLINE=96241580; PubMed=8673113; | | | |
| RA | Weisner P.N., Dailey T.A., Hitt R.J., Ziman M., Corriganl A.V., | | | |
| RA | Roberts A.G., Weisner D.M., Kirsch R.E., Dailey H.A.; | | | |
| RT | "A R5W mutation in human protoporphyrinogen oxidase results in | | | |
| RT | decreased enzyme activity and is prevalent in South Africans with | | | |
| RL | variegate porphyria."; | | | |
| RL | Nat. Genet. 13:95-97(1996). | | | |
| RN | [5] | | | |
| RP | VARIANT VP CYS-152. | | | |
| RX | MEDLINE=98434271; PubMed=9763307; | | | |
| RA | Frank J., Poh-Fitpatrick M.B., King L.E. Jr., Christiano A.M.; | | | |
| RT | The genetic basis of 'Scarsdale gourmet diet' variegate porphyria: a | | | |
| RT | tandem mutation in the protoporphyrinogen oxidase gene."; | | | |
| Arch. Dermatol. Res. 290:441-445(1998). | | | | |
| -1- | FUNCTIO: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYNOGEN | | | |
| CC | IX TO FORM PROTOPORPHYN IX. | | | |
| CC | -1- CATALYTIC ACTIVITY: PROTOPORPHRYNOGEN-IX + O(2) = PROTOPORPHYRN- | | | |
| CC | IX + H(2)O(2). | | | |
| CC | -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER. | | | |
| CC | -1- PATHWAY: PENULTIMATE STEP INHEME AND PORPHYRN BIOSYNTHESIS. | | | |
| CC | -1- SUBUNIT: HOMODIMER. | | | |
| CC | -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE | | | |
| CC | WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY). | | | |
| CC | -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG, | | | |
| CC | LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS. | | | |
| CC | -1- DISEASE: DEFECTS IN PPOX ARE THE CAUSE OF PORPHYRIA VARIEGATA | | | |
| CC | (VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR | | | |
| CC | HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE | | | |
| CC | INTERMITTENT PORPHYRIA. | | | |

| | |
|----|---|
| CC | -I- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY. |
| CC | ----- |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL Outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use by non-profit institutions as long as its content is in no way |
| CC | modified and this statement is not removed. Usage by and for commercial |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-ch). |
| CC | -- send an email to license@isb-sib.ch . |
| CC | ----- |
| DR | EMBL; D38537; BAA07538.1; -; |
| DR | EMBL; U26446; AAA67690.1; -; |
| DR | MIM; 600923; -; |
| DR | MIM; 176200; -; |
| KW | Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase; |
| KW | FAD; Mitochondrion; Disease mutation; Polymorphism. |
| FT | NP_BIND |
| FT | 9 |
| FT | 14 |
| FT | RAD (POTENTIAL). |
| FT | R -> W (IN VP). |
| FT | /FTID=VAR_003686. |
| FT | R -> C (IN VP). |
| FT | /FTID=VAR_003687. |
| FT | VARIANT |
| FT | 152 |
| FT | 152 |
| FT | R -> C. |
| FT | /FTID=VAR_003688. |
| FT | G -> R (IN VP). |
| FT | /FTID=VAR_003689. |
| FT | R -> H. |
| FT | /FTID=VAR_003690. |
| FT | VARIANT |
| FT | 168 |
| FT | 168 |
| FT | G -> R (IN VP). |
| FT | /FTID=VAR_003691. |
| FT | R -> H. |
| FT | /FTID=VAR_003692. |
| FT | VARIANT |
| FT | 232 |
| FT | 232 |
| FT | G -> R (IN VP). |
| FT | /FTID=VAR_003693. |
| FT | R -> H. |
| FT | /FTID=VAR_003694. |
| FT | VARIANT |
| FT | 304 |
| FT | 304 |
| FT | R -> H. |
| FT | /FTID=VAR_003695. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003696. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003697. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003698. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003699. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003700. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003701. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003702. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003703. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003704. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003705. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003706. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003707. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003708. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003709. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003710. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003711. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003712. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003713. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003714. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003715. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003716. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003717. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003718. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003719. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003720. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003721. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003722. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003723. |
| FT | VARIANT |
| FT | 507 |
| FT | 507 |
| FT | R -> H. |
| FT | /FTID=VAR_003724. |
| | |

| | | | | | |
|-----------------------|--------------|--|----------------|------------|---------|
| Query Match | 12.8% | Score 363.5 | DB 1 | Length 477 | |
| Best Local Similarity | 25.5% | Pred. No. 5e-20 | | | |
| Matches 138 | Conservative | 83 | Mismatches 170 | Indels 151 | Gaps 23 |
| QY | 75 | VIVGAGISGLCIAQVMSANY-----PMLAVTARDRAGCNITTV- RDGYLWEEGPN | 125 | | |
| | | | | | |
| Db | 6 | VYLGCGISGL-----AASYHLSRAPCPKPVYLVESSEHSGMINSQGNPAIFELGPR | 59 | | |
| QY | 126 | SFOEEDPM-----LTMAVDCGLKDDL--VLGDNPAP--RFLIMKGLRPVPSKLTDLPEFD | 177 | | |
| | | | | | |
| Db | 60 | GIRPAGAGARTILLVSEIGDSEVLPVGRDHPAQNRFYLVGALHMLPTGLRCL--- | 115 | | |
| QY | 178 | IMSIGKLRAGFGPIRSPSP-----GHE--ESVQFVRNNGEYF | 218 | | |
| | | | | | |
| Db | 116 | -----LRSPPEFKPLPWAGLRKELTKRGKPEDETIVSFAQRRLGPEYA | 159 | | |
| QY | 219 | RLIIPFCSGYVVDPSKLSMKAFAKVKWLTETGSGIIGTFKAIKERSSTPKAPDRP | 278 | | |
| | | | | | |
| Db | 160 | SLAMSDLCRGVAFAGNSRELSTIRSCPSLFQAEOTHRSLTGLLGA-----GR | 207 | | |
| QY | 279 | LPRKPGQIV-----GSFRKGLNMLPDAISARLSK-----LKLST----- | 312 | | |
| | | | | | |
| Db | 208 | TPQPSALIRQALABRMSOWSLRGLEMLPOLLETHLTSKGVSVLRGQVCGLSIQAEGR | 267 | | |
| QY | 313 | WKLSSITKSEKGGYHLTETPEGVVSSLOSRSIVMTVPSSVASNILRPLSVAAADALSNEY | 372 | | |
| | | | | | |
| Db | 268 | WKV--SLRDS-----SLDADHYISAIPASVISELLPAEAPLALASLMT | 310 | | |
| QY | 373 | YRPVCAVITISFQEAIRDBRVLVDGELKFGQGLHPTQGVETLIGTYSLPPNR--ARKG | 430 | | |
| | | | | | |
| Db | 311 | AVSAVAVMLQYOGAHL-----PVQGFGLHVPSESDPGVGIYDVAFPEQDSPPG | 362 | | |
| QY | 431 | RYLLANTIGGAKNNEILLKTESOLVEVVDRLRKMLKPKRQODLV--VGRVRVP---- | 483 | | |
| | | | | | |
| Db | 363 | LRAVVM--LGG-----SWLQTLNASCYLSOELTQOQRQEAALPOLGLIKEPSSHCL | 411 | | |
| QY | 484 | -----QAIPQELVGHLDLTSTAKAAMNNGLEGLGNGYVSGVALGRGVEGAYEVASEV | 538 | | |
| | | | | | |
| Db | 412 | VHLHKNCIPQYTLTGWQGLSARQGLTAHRLPLTLAGASY--EGVAVNDCIESGRRAAASV | 470 | | |
| QY | 539 | TG 540 | | | |
| | | | | | |
| Db | 471 | LG 472 | | | |

RESULT 8
PPOX_MOUSE STANDARD; PRT: 477 AA.
ID PPOX_MOUSE P51175; P97344;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN PPOX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96132666; PubMed=8554330;
RT Dailley T.A., Dailley H.A., Weissner P., Prasad A.R.;
RT "Cloning, sequence, and expression of mouse protoporphyrinogen
RT oxidase.";
RT Arch. Biochem. Biophys. 324:379-384(1995).
RL [2]
RN SEQUENCE FROM N.A.
RC TISSUE=Erythrocytes;
RX MEDLINE=95331315; PubMed=7607249;
RA Takekani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,
RA Nishimura K., Inokuchi H.;
RT "Induction of terminal enzymes for heme biosynthesis during
RT differentiation of mouse erythrocytic cells.";
RL Eur. J. Biochem. 230:760-765(1995).
RN [3]
RP SUBCELLULAR LOCATION.
RC TISSUE=Liver;
RX MEDLINE=88153682; PubMed=3346226;
RA Ferreira G.C., Andrew T.L., Kerr S.W., Dailley H.A.;
RT "Organization of the terminal two enzymes of the heme biosynthetic
RT pathway. Orientation of protoporphyrinogen oxidase and evidence for a
RT membrane complex.";
RL J. Biol. Chem. 263:3835-3839(1988).
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC -1- PATHWAY: PENDING STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
CC -1- INDUCTION: DURING ERYTHROID DIFFERENTIATION.
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U25114; AAA96003.1; -
CC EMBL: D45185; BAA08126.1; -
CC DR MGD: MG1:104968; PPOX.
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
KW Flavooprotein; FAD; Mitochondrion.
FT NP_BIND 9 14 FAD (POTENTIAL).
FT CONFLICT 64 64 A -> T (IN REF. 2).
FT CONFLICT 66 66 A -> P (IN REF. 2).
FT CONFLICT 108 108 L -> S (IN REF. 2).
FT CONFLICT 427 427 W -> C (IN REF. 2).
SQ SEQUENCE 477 AA; 50870 MW; 8CFB48120728DE6F CRC64;

Query Match 12.7%; Score 362.5; DB 1; Length 477;
Best Local Similarity 24.8%; Pred. No. 5.9e-20;
Matches 133; Conservative 76; Mismatches 188; Indels 139; Gaps 19;
QY 75 VIVGAGISGLCIAQ--VSNANPNMVTIARDRAGNTTYE-RDGYIWEBCGNSFQPSD 131
DB 6 IVLGGISGLIAASYHLIRGSPKYLIVEGSKRLGWSIRSGSIAFELGPRGIRPAG 65
QY 132 PM---LTMAVCCGLKDDL-VLGDPNAP--REFVLMKRLRPVPSKLDLPFFDLMSIRG 183
DB 66 ALGARILLVSELGLSEVLPVRGDHPAQNRELVGGTILHPLPSGLNGL----- 115
QY 184 KLRAGFGPIGLRSPSP-----GHE--ESYEQFYRMNMGVEFERLIEP 224
DB 116 -----LRRPSPFKPLRWAGLRELLKRGKEPEPETHSFQRRLGFEVANSIANDS 165
QY 225 FCSGYVVDPSKLSKAAFGKWKLEFSGSTIGTFEAIKERSSTPRAPDPLPKPG 284
DB 166 LCRGVAGNSRELIRSCFPPLFQAEOYHRSILGLIGAGOS-----POPDS 213
QY 285 QTV-----GSEFRGLRMLPDAISARLGSK-----TKLS---WKLSI 318
DB 214 SLIRQARRRMSQWSLRGLGVLPQALNHLASKGVTVLSCQPVCGSLQDPGRKXVSL- 272
QY 319 TSEKGGYHLIETPEGVVSLQSRISVTVPYSVSNILRPLSYAADALSNFYYPVGA 378
DB 273 -----GDSLEADHIIISAIASELSKILLPRAAPLAILITIRAVSYAV 316
QY 379 VTIISFQEARIRDERLVDEBLKFGQLHPTQGVETLGTIYSSLPNR--APKGRVILL 435
DB 317 VMLQFRGACL-----PVQGFHILVPSSDEPVLGIYVDSVAFEPEDGNPPLRYVM 368
QY 436 NYIGAKNPILLKTESOLVEYVDRDLRKLKPKAOD-----PLVGVRRWPD 484
DB 369 --LGG-----YWLQKLAAGHQLSPBELFQQAQQAALATQGLKPPSHCLVHLKKN 417
QY 485 ALPQFLVGHLDLTSTAKAAMDNGLEGLGNYVYVAGLCRCYGAEVASEYTG 540
DB 418 CIPQYTIQHWQKLDISAMQFLTAQRLLPLTLGASV-EGVAVNDCIESGQAQAVAVLG 472
RESULT 9
YAM7_SCHPO STANDARD; PRT: 490 AA.
ID YAM7_SCHPO Q10062;
AC 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOPHETICAL 53.5 KDA PROTEIN CLP5.07C IN CHROMOSOME 1.
GN SPAC1P5.07C.
OS Schizosaccharomyces pombe (Fission Yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetes; Schizosaccharomycetaceae;
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=972;
RA Gentles S., Odell C., Churcher C.M., Barrell B.G., Rajandream M.A.,
RA Walsh S.V.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) AND TO
CC YEAST YER014W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z68136; CAA92235.1; -
CC DR

KW Hypothetical protein
SQ SEQUENCE 490 AA; 53461 MW; CB679F6A437547AA CRC64;

Query Match 10.6%; Score 302.5; DB 1; Length 490;
Best Local Similarity 22.9%; Pred. No. 2.1e-15;
Matches 120; Conservative 103; Mismatches 213; Indels 87; Gaps 20;

QY 72 LDCVVGAGISGLCIAQVMSANYPNLMTV--EARDRAGNITVE-----RDGYLMEB 122
D 1 MSIAICGGIAGLSTAFYLRIPIKCTIDLYEKGPRILGWLQSVKIPCADSPYGLVLFED 60
QY 123 GPNSEFQSD-----PMLTMAVDCGILKMDLYIGDPNAP---REFLMGKLRPVSKLTLDLP 174
D 61 GPRLRLRAGVAGLANLDLISKLGIEDKLRISSNSAKRRITYYDRLEIIS-----114
QY 175 FFDLMSIPGKLIRAGFGPIGLRSP-----PGHESEVEQFVRNMLGGEVEERL 221
D 115 -----SILGSIKSIMP-ALRPMPLAMLEPFKRSKSDSDSEVSGPMRRRGKNTDRY 168
QY 222 IEPFGSGVYVGDPSKLSMKAA-EGKWKLDETGSGIIGFFAKTER--SSTPKAPRDP 277
D 169 MSAMINGIYAGDLNDLSMSSMFGFLAKIKKYGNTILGIRALLRLAREILSPAELKAA 228
D 278 RLPRP-----KGTVSGFRKGLRMLPDALSARLSKILKLSITKSEKGYHL 328
D 229 LLAEPKTAELSNMSKTSIMAFKEGIEITITLSTADELKKPNVKIHLNKRFT-----L 282
QY 329 TYETPEGVSLQSRSIYMTVPVSYASNILRPLSVAADALSINFYPPVPGAVTISYQEA 388
D 283 VPHKQGLVLDVNGGAYEYVAFANSRLNLEMLISCPKMEPTSSVY---VVNYVY----- 333
QY 389 RDELVLVGEELKGFQOLPR---TQGVETLGTIYSSSL-FENRPAKGVILLNTYIGA--- 441
D 334 KDPVNL--PIRGFLLIPSCPTPNPNPHVLGIVDSEQNENSGSKYVVM-----GGSA 387
QY 442 KNPFIILKTESQLEVEVDRDLRKLML-TPKRAODPLVGVVWVPOALPQFLVGHLDLTSTA 500
D 388 KNTSLIPNPEAVNNMLKALQHTLKTSK---PTLLNATLQONCIPQYVAVGHODNLNSL 444
QY 501 KAANNNGLEGLGNYVSGVALGRCVEGAYEAVESVTFPLS 543
D 445 KSWIEKNMGGRILLTGSWYNGVSGIDCIMGHSPARLASLMN 487
RESULT 10
PPOX_MYCTU STANDARD; PRT; 426 AA.
AC 053230;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
GN HEMV OR RV2677C OR MYV010.01C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID-1773;
RA [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE=98295987; Pubmed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).

CC - FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
CC IX TO FORM PROTOPORPHYRIN IX.
CC - CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) - PROTOPORPHYRIN-
CC IX + H(2)O(2).
CC - COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
CC - PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
CC - SUBUNIT: HOMODIMER (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
CC - CAUTION: SEEMS SHORTER IN THE C-TERMINAL THAN CORRESPONDING HEMV
CC IN OTHER BACTERIA.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AL021186; CA16020.1; -
DK Tuberculin; RV2677C; -
KW Protophylin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
KW Heme biosynthesis.
FT NP_BIND 10 15 FAD (POTENTIAL).
SQ SEQUENCE 426 AA; 44382 MW; 221952B3977D380 CRC64;

Query Match 7.8%; Score 221; DB 1; Length 426;
Best Local Similarity 22.8%; Pred. No. 2.5e-09;
Matches 108; Conservative 74; Mismatches 210; Indels 82; Gaps 17;

QY 76 IYGAIGSGICLA-QVMSANYPNLMTV--EARDRAGNITVEERDYLTEEGSPNSFQSDP 132
D 8 VVGGSIGSLTAYRLRQAVGDATTITFEPRDLRGVLRTEHIGGQPMDLAEAFVLRP 67
QY 133 -MLTMAVDCGILKMDLYIGDPNAPFVLMKGLRPVPSKLTLDLPFFDLMSIPGKLIRAGCP 191
D 68 EMPALLAEGL-SDROLASTGARPLYSQORLHPDP-----TVVGISPSAGSMAGL 119
QY 192 I-----GLRPS--PGHESEVEQFVRNMLGGEVEERLIEPFGSGVYVGDPSKLS 238
D 120 VDDATLARIADAARPEFTWQVSDPAVADLVADRGDOVAVSPLISGYAGSAATIG 179
QY 239 MKAAFGKVMKLEETGSGIIGTFKAIKERSYPAKAPRDLRPGKQTVSGFRKLRLMP 298
D 180 LRRAAPSVAAALDRGA-----TSYTDVAVRQALPFGSGGPGFEGALDGGYVLL 226
QY 299 DAISARLSKLSKLSITKSEKGYHLIYETPEGVSYLQSRSIYMTVPVSYASNILR 358
D 227 DGLVRR---SHVHWRAVYQDLER-GWVLREDETG---RWQADAVIILAVAPRLARLVD 278
QY 359 PLSTVAADALSINFYPPVPGAVTISYQEAIRDERLVDELKGFQGLHPTQGVETLG--- 415
D 279 GLAPRTHAARQIVSASSAVVALAVPG-----GTAFFHSCSVLYVAGDES 322
QY 416 -----TIYSSSLFPRAPKGRVILLNTYIGAKNPE-ILSKTESQLEVEVDRDLRKLMLIK 468
D 323 PHAKATITLSSRW---GGRGDVALLRLSFGFRGDEPALTLASDDQLLMAAADL--VTVF 376
QY 469 PRAODPLVGVVWVPOALPQFLVGHLDLTSTA KAANNNGLEGLGNYVSGV 522
D 377 GVAVDVAVRVRWIEAMPQYGPCHADVAELRAGLPT---LAVAGSYLDGI 426
RESULT 11
PPOX_MYCTU STANDARD; PRT; 451 AA.
AC 050008;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).

GN HEMV.
 OS Mycobacterium leprae.
 OC Bacteria; Filimicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.;
 RL Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).
 CC -1- PATHWAY: PENTULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U15181; AAA62958.1; -;
 KW Protophyrin biosynthesis; Oxidoreductase; Flavoprotein; FAD;
 FT NP_BIND 10 15 FAD (POTENTIAL).
 SQ SEQUENCE 451 AA; 46580 MW; D76BE1655CA2056 CRC64;

Query Match 7.6%; Score 216.5; DB 1; Length 451;
 Best Local Similarity 20.6%; Pred. No. 5.8e-09;
 Matches 105; Conservative 79; Mismatches 193; Indels 133; Gaps 18;

QY 76 IVGAGISGLCTIA---QWSANPYNLMTEARDRAGKNTTVERDGYLWEGEENSF---QP 129
 DB 8 VVGGISGLTAAAYRLRAVATGDDVAITLFPDGRIGVLRCEVGGOGPMDDGAFAFLRP 67
 QY 130 SDPMLTMAVDCGLND-----DLVLSGPNAPRYLWMLKGLRP 166
 DB 68 EVPAL--LAELGLSERQATTDARPLIYSQORLSLPPDTVAG-----I 109
 QY 167 PSKLTLPFFDLMSIPKLR---AGFGPIGLRPS--PGHESEVEQFVRNLGGEYER 220
 DB 110 PSSAT-----SVAGLVDAVATVARIQAENVRPLSWEPSPDPMALVADREGEQAVAR 161
 QY 221 LIEFPCGGVYVYDPSKLSMKAAPGKWKLETTGGSTIGTFKAIKERSSTPKAPRDRPLP 280
 DB 162 LVDPILGGVYVAGSAATTGLRAGASVAAALDCGA-----TSLMEAVRQGLPP 208
 QY 281 KPKQGVTSFRRKGLRMLPDALISARLGSKLKSLKSLSTFSEKNG--GYHLTYETPECVS 338
 DB 209 VAAAPVPGALDGGVQVYLDELVRR-----SRLOVAAATVVGIDKGTGWTLVDDG---AC 261
 QY 339 LOSISIVYVPSYVANSILRPLVAAADALSNEYPPVGAVTISYPOEARLDERLYVDGEL 398
 DB 262 WSAAGVILAVPAPRLVRLDQOIAFRVAAASRIYASASSAVVALSVPRDTTFPQN--SGVL 319
 QY 399 KGFGQ-----LHPTQGVETLGTIYSSSLFPNAPKRVLLANTIG-----A 441
 DB 320 VASGERLRAAVVILSSKKWGLQ-----GDTOLVRLSFGKFPDQVASTA 362
 QY 442 KNPILSKTESOLVEVVDRLRKLKPKAODPLVGVVRWPAOIFOLVGHLDITLSTAK 501
 DB 363 SDELLAMAVSDILAAYVDVY-----DPYDVCYQKMDIAMQYRGHADVAAEVR 412
 QY 502 AAMNDNGLBELFLGNGVSVSVALGRCEYGA 531
 DB 413 AGLPPT---LVVAGSHMDIGVAPACISAA 438

RESULT 12
 ID PROX_YEAST STANDARD; PRT; 539 AA.
 AC P40012;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4) (PPO).
 GN HEM14 OR YER014W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96224138; PubMed=8621563;
 RA Camadro J.-M., Labbe P.;
 RT Cloning and characterization of the yeast HEM14 gene coding for
 RT protoporphyrinogen oxidase, the molecular target of diphenyl
 RT ether-type herbicides.";
 RL J. Biol. Chem. 271:9120-9128(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berro A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunnicke-Smith S.,
 RA Hyman R., Kayser A., Komp C., Lashkari D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Namath A., Norgren S., Oefner P., Oh C.,
 RA Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SUBCELLULAR LOCATION.
 RC STRAIN=ATCC 25657 / D273-10B;
 RX MEDLINE=95096046; PubMed=7798202;
 RA Camadro J.-M., Thome F., Brouillet N., Labbe P.;
 RT Purification and properties of protoporphyrinogen oxidase from the
 RT yeast Saccharomyces cerevisiae. Mitochondrial location and evidence
 RT for a precursor form of the protein.";
 RL J. Biol. Chem. 269:32085-32091(1994).
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN
 CC IX TO FORM PROTOPORPHYRIN IX. OPTIMUM ACTIVITY IS OBTAINED AT PH
 CC 7.2.
 CC -1- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) = PROTOPORPHYRIN-
 CC IX + H(2)O(2).
 CC -1- COFACTOR: FAD; CONTAINS 1 MOLE OF FAD PER MOLE OF ENZYME.
 CC -1- PATHWAY: PENTULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z71381; CAA95981.1; -;
 DR EMBL; U18778; AAB64547.1; -;
 DR SGD; S0000816; HEM14.
 KW Protophyrin biosynthesis; Heme biosynthesis; Oxidoreductase;
 KW Flavoprotein; FAD; Mitochondrion.
 FT NP_BIND 18 23 FAD (POTENTIAL).
 FT MUTAGEN 422 422 L->P: IN HEM14-1; LOSS OF ACTIVITY.
 FT MUTAGEN 424 424 K->E: IN HEM14-1; LOSS OF ACTIVITY.
 SQ SEQUENCE 539 AA; 59703 MW; 5ECBBD1C033BA5B1 CRC64;


```

RT      "A novel, soluble form of phytoene desaturase from Narcissus
RT      pseudonarcissus chromoplasts is Hsp70-complexed and competent for
RT      flavinylation, membrane association and enzymatic activation."
RT      Plant J. 9:601-612(1996).
RN      [3]
RP      SUBCELLULAR LOCATION.
RX      MEDLINE=97433276; PubMed=9288918;
RA      Bonk M., Hoffmann B., von Lintig J., Schledz M., Al-Babli S.,
RA      Hobeika E., Kleinig H., Beyer P.;
RT      Chloroplast import of four carotenoid biosynthetic enzymes in vitro
RT      reveals differential fates prior to membrane binding and oligomeric
RT      assembly".
RT      Eur. J. Biochem. 247:942-950(1997).
CC      -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
CC      INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
CC      DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
CC      -1- COFACTOR: FAD (PROBABLE).
CC      -1- PATHWAY: CAROTENOID BIOSYNTHESIS.
CC      -1- SUBCELLULAR LOCATION: CHLOROPLAST (CHROMOPLAST). EXISTS AS AN
CC      INACTIVE SOLUBLE FORM AND AN ACTIVE MEMBRANE-BOUND FORM.
CC      -1- TISSUE SPECIFICITY: EXPRESSED MORE STRONGLY IN FLOWERS THAN IN
CC      LEAVES.
CC      -1- DEVELOPMENTAL STAGE: EXPRESSION INCREASES DURING FLOWER
CC      DEVELOPMENT.
CC      -1- SIMILARITY: TO OTHER PLANTS OR CYANOBACTERIAL PHYTOENE
CC      DESATURASES.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: X78815; CA55392.1; -
DR      Mendel: J3184; Harps; P051; J3184.
DR      InterPro: IPR002037; -
DR      Pfam: PF01593; Amino-oxidase; 1.
DR      Carotenoid biosynthesis; Oxidoreductase; FAD; Chloroplast; Membrane;
KM      Transit peptide.
FT      TRANSIT 1 91 CHLOROPLAST (POTENTIAL).
FT      CHAIN 92 570 PHYTOENE DEHYDROGENASE.
FT      NP_BIND 104 120 FAD (ADP PART) (BY SIMILARITY).
SQ      SEQUENCE 570 AA; 63790 MW; 654F569F5B83BE77 CRC64;

```

Query Match 4.9%; Score 139; DB 1; Length 570;
 Best Local Similarity 21.1%; Pred. No. 0.0058;
 Matches 111; Conservative 92; Mismatches 178; Indels 146; Gaps 24;

```

DB      351 LGNGNITGDAVY-----VAA-----PVDILKLIPQEWREIPYFKLD-K 390
OY      398 LKGFGLHPRTOGVETLIGTYSSLF-----PNRAPGRVLLN 436
DB      391 LAGVPYINAHINWFDRKLTNYDHLFTSPSPILSYAADMVCKREYDPNS-----MELV 446
OY      437 YIGAKNPETLSTESQOLVEVVDRLRMKLKPAQDEPLVGVFWPQAIPOPLVGLDT 496
DB      447 F---AAAEWISRSSELIETKMKELAKLPDELAAD-----QSKAKIKYHV-- 491
OY      497 LSTAKA-----MNDNGLEGLGNGVSGVALRCVEGA 531
DB      492 VKTPRSYKTIPOCEPCRLQSPLEGFLAODITNOKYLA-SMEGA 537

```

RESULT 15
 AOF_ONCMY STANDARD; PRT; 522 AA.
 AC P49253;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE AMINE OXIDASE [FLAVIN-CONTAINING] (EC 1.4.3.4) (MONOAMINE OXIDASE) (MAO).
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=95107295; PubMed=7808446;
 RA Chen K., Wu H.-F., Grimsby J., Shih J.C.;
 RT Cloning of a novel, monamine oxidase cDNA from trout liver.";
 RL Mol. Pharmacol. 46:1226-1233(1994).
 RN [2]
 RP IDENTIFICATION OF PROBABLE FRAMESHIFT.
 RA Balroch A.;
 RL Unpublished observations (JUL-1998).
 CC -1- FUNCTION: OXIDIZES BOTH 5-HYDROXYTRYPTAMINE (5-HT) AND BETA-
 CC PHENYLETHYLAMINE (PEA)
 CC -1- CATALYTIC ACTIVITY: RCH(2)NH(2) + H(2)O + O(2) = RCHO + NH(3) +
 CC H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL OUTER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
 CC -1- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN IN THAT, DUE TO A
 CC PROBABLE FRAMESHIFT, IT ENDS IN POSITION 499 AND LACKS THE C-
 CC TERMINAL REGION FOUND IN OTHER MEMBERS OF THIS FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L37878; AAA64302.1; ALT_FRAME.
 DR InterPro: IPR001613; -
 DR InterPro: IPR002937; -
 DR Pfam: PF01593; Amino-oxidase; 1.
 DR PRINTS: PR00757; AMINEOXIDASEF.
 KM Oxidoreductase, Flavoprotein; FAD; Transmembrane; Mitochondrion.
 FT NP_BIND 7 63
 FT BINDING 399 399 FAD (ADP PART) (POTENTIAL).
 FT TRANSMEM 493 513 FAD (BY SIMILARITY).
 SQ SEQUENCE 522 AA; 58937 MW; F82D6D15364D646D CRC64;

Query Match 4.7%; Score 134; DB 1; Length 522;

Best Local Similarity 22.0%; Pred. No. 0.012;
Matches 67; Conservative 48; Mismatches 119; Indels 70; Gaps 12;

```
QY 73 DCYIVAGISGLCIAQVMSANPNLMTVEARDRAGNITTVTERDGYLME-----GPSN 126
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 8 DYIVIGGISGLSAARLKEKGLSPVLEARDRVGGRTFTVONEQTKYVDLGGAYVGPT- 66
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 127 FQSPDMLTMAVDCGLKDDLVLDPNAPRFVLMKGLRPVPSKLTDL--PF----- 175
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 67 ---QNRILRLAKRCGYKTIKVNNEERLVHY--KGRSYPFKGSFPPMNNPFALMDYNNLM 121
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 176 --FDLM--SIPGKLRAFGPIGLRPSPPGHESVEQFVRRNLGGEV-----FERLIEPFC 226
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 RKMDENGSEIIPRE-----APWKAPHAEMWDKMTMKQLFDKICWTSSARRFATLFFV 171
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 227 SGYYVGDPSKLSMKAFFGKVMKLEETGGSF-IGTFKAIKERSSTPKAPRDPRLPKPGQ 285
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 172 NVNVTSEPHVS--ALMFLMTYVKQCGGTMRIFSTINGQER----- 210
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 286 TVGSFRKGLMLPDATISARLGSKLKLSWKLSITKSEKGGYHLTYETPEGVVSLOSRSIV 345
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 211 ---KFLGGSSQISECAKELGERVKME---SPYKIDQTDWVEVETLNKETYKAKAYIV 264
      |:::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 346 MIVP 349
      | | |
Db 265 ATPP 268
```

Search completed: July 3, 2001, 10:43:19
Job time: 612 sec

OM of: US-09-508-418-2 to: EST:* out_format: pfs

Date: Jul 3, 2001 11:06 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

MODEL=frame+;p2n.model -DEV=xlp
-O=/cgn2_1/uspro.spool/US09508418/runcat-03072001.100016.24427/app_query.fasta-1.613
-DB=EST -QPM=fastap -SUFFIX=first -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOCPCL=0.000 -LOAPEXT=0.000 -GAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-DELEXPT=7.000 -XGAPEXT=10.000 -XGAPEXT=0.500 -DELOP=6.000
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09508418.@CGN1_13324 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-508-418-2
Query length: 548
Database: EST:*
Database sequences: 10228115
Database length: 431459454
Search time (sec): 1358.990000

score_list:

| Sequence | Strd Orig | ZScore | EScore | len | Documentation |
|--------------------|-----------|---------|---------|-----|----------------------------------|
| gb_est73.BE343078 | + 891.50 | 1639.28 | 3.4e-82 | 593 | BE343078 EST395922 potato stolon |
| gb_est50.AW695211 | + 838.00 | 1538.73 | 1.4e-76 | 640 | AW695211 NE092606S1F1051 Dve1 |
| gb_est99.BG320666 | + 822.50 | 1507.46 | 7.6e-75 | 780 | BG320666 Zm04.06e01.R Zm04.AAPC |
| gb_est99.BG322999 | + 797.00 | 1463.39 | 2.2e-72 | 580 | BG322999 sbp93f10.Y1 Gm-cl040.G |
| gb_est74.BE45801 | + 778.00 | 1432.22 | 2.0e-70 | 580 | BE45801 HVSBE90015D05f Hordeum |
| gb_est39.AV38014 | + 739.00 | 1352.64 | 2.3e-66 | 596 | AV38014 AV38014 Arabidopsis t |
| gb_est30.BF654024 | + 679.00 | 1255.64 | 7.1e-61 | 510 | BF654024 HVSBE90015D05f Hordeum |
| gb_est42.AW119581 | + 624.00 | 1236.44 | 4.1e-55 | 498 | AW119581 sda8d01.Y1 Gm-cl028.GJ |
| gb_est69.BE021666 | + 629.00 | 1153.11 | 4.1e-55 | 480 | BE021666 sm61b05.Y1 Gm-cl028.GJ |
| gb_est39.BG353389 | + 621.00 | 1130.24 | 2.8e-54 | 479 | BG353389 947043H05.Y2 947 - 2 W |
| gb_est39.AV540305 | + 605.00 | 1107.47 | 1.4e-52 | 519 | AV540305 AV540305 Arabidopsis t |
| gb_est42.AW174851 | + 593.00 | 1086.68 | 2.1e-51 | 457 | AW174851 sb54d01.Y1 Gm-cl016.GJ |
| gb_est97.BG133906 | + 589.00 | 1079.70 | 5.1e-51 | 440 | BG133906 EST466798 tomato crown |
| gb_est73.BE336089 | + 552.50 | 1008.16 | 4.9e-47 | 590 | BE336089 NE084G10S1F1081 Dve1 |
| gb_est51.AW755783 | + 543.00 | 993.75 | 3.1e-46 | 451 | AW755783 s109c11.Y1 Gm-cl036.GJ |
| gb_est70.BE129089 | + 534.00 | 974.95 | 3.5e-45 | 533 | BE129089 894020B03.Y1 C. reinha |
| gb_est42.AW024400 | + 525.00 | 961.28 | 2.0e-44 | 414 | AW024400 s14e09.Y1 Gm-cl027.GJ |
| gb_est42.AW091632 | + 492.00 | 902.35 | 3.8e-41 | 337 | AW091632 EST84908 tomato mixec |
| gb_est39.AV51750 | + 491.00 | 893.98 | 1.1e-40 | 574 | AV51750 AV51750 Arabidopsis t |
| gb_est27.A1997032 | + 460.50 | 840.87 | 1.0e-37 | 425 | A1997032 T0153155 A. thaliana, |
| gb_est72.BE229685 | + 431.00 | 783.30 | 1.7e-34 | 530 | BE229685 98SNO073 Rice Immature |
| gb_est101.BG462577 | + 419.00 | 761.30 | 2.8e-33 | 513 | BG462577 947058C08.X1 947 - 2 W |
| gb_est39.AV50160 | + 415.00 | 756.47 | 4.5e-33 | 446 | AV50160 AV50160 Arabidopsis t |
| gb_est11.AA712377 | + 415.00 | 756.47 | 5.1e-33 | 414 | AA712377 32105 Lambda-PR2 Arab |
| gb_est39.AV540703 | + 408.00 | 744.26 | 2.5e-32 | 387 | AV540703 AV540703 Arabidopsis t |
| gb_est27.A1991783 | + 398.00 | 723.11 | 3.7e-31 | 476 | A1991783 618035B09.X1 618 - Ind |
| gb_est47.AW433355 | + 394.00 | 714.99 | 1.0e-30 | 503 | AW433355 660040C08.Y1 660 - Mix |
| gb_est46.AW428647 | + 388.00 | 702.08 | 1.5e-30 | 580 | AW428647 660040C08.X2 660 - Mix |
| gb_est16.BE601207 | + 365.00 | 659.67 | 1.3e-27 | 561 | BE601207 P11_90.C06.G1 A002 Pat |
| gb_est31.AU175016 | + 359.00 | 653.23 | 3.3e-27 | 413 | AU175016 AU175016 Rice cDNA frg |
| gb_est101.BG462578 | + 357.50 | 653.38 | 2.8e-27 | 299 | BG462578 947058C08.Y1 947 - 2 W |
| gb_est53.AW929479 | + 351.00 | 639.50 | 1.7e-26 | 346 | AW929479 EST838267 tomato flowe |
| gb_est31.AU175015 | + 276.00 | 495.78 | 1.7e-18 | 483 | AU175015 Rice cDNA frg |
| gb_est51.AW781134 | + 272.00 | 488.45 | 3.8e-18 | 441 | AW781134 s189e02.Y1 Gm-cl037.GJ |
| gb_est43.AW313289 | + 257.00 | 460.90 | 1.5e-16 | 464 | AW313289 660023B03.X1 660 - Mix |
| gb_est49.AW615911 | + 253.50 | 445.51 | 9.0e-16 | 587 | AW615911 EST325409 tomato flowe |
| gb_est100.BG453337 | + 251.50 | 446.34 | 9.6e-16 | 661 | BG453337 NF096C05L1F1035 Dve1 |
| gb_est109.T43573 | + 246.00 | 440.26 | 2.1e-15 | 470 | T43573 6836 Lambda-PR2 Arabid |
| gb_gss19.AQ954144 | + 220.00 | 393.08 | 8.1e-13 | 425 | AQ954144 nbep0054K15R CUGI Rice |
| gb_est90.BF598346 | + 215.50 | 381.03 | 4.2e-12 | 574 | BF598346 svt16e03.Y1 Gm-cl056.GJ |
| gb_est50.AW655910 | + 215.50 | 380.66 | 4.4e-12 | 592 | AW655910 NF099RH7.S1F1063 Dve1 |

gb_gss17.AQ858291 - 214.00 381.65 3.9e-12 434 | AQ858291 nbep0012122f CUGI R
gb_est51.AW755660 + 213.50 376.97 7.0e-12 590 | AW755660 s106e04.Y1 Gm-cl036
em_est15.AW726965 + 207.50 361.03 5.4e-11 872 | AW726965 GA_Ea0023F07 Gossy
gb_est100.BG444110 + 207.50 361.03 5.4e-11 872 | BG444110 GA_Ea0023F07 Gossy

seq_name: gb_est73.BE343078

seq_documentation_block: 593 bp mRNA EST 17-JUL-2000
LOCUS BE343078
DEFINITION EST395922 potato stolon, Cornell University Solanum tuberosum cDNA
clone cSTA2218, mRNA sequence.
ACCESSION BE343078
VERSION BE343078.1 GI:9252610
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 593)

REFERENCE
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Uiterback,T., Bowman,C.L., Doan
,B., Bougri,O., Bell,C.R., Konning,C.M., Tanksley,S.D. and Baker
,B.
Generation of ESTs from potato swelling stolons
Unpublished (1999)
Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com

FEATURES
source
Location/Qualifiers
1..593
/organism="Solanum tuberosum"
/cultivar="hintje"
/db_xref="taxon:4113"
/clone="cSTA2218"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2: XhoI; RNA was supplied by Christian Bachem & Beatrix Horvath(laboratory of Plant Breeding, Dept of Plant Sciences, Wageningen University, The Netherlands). Total RNA was isolated from developing axillary buds of potato nodal stem cuttings cultured on medium for the introduction of tuber formation as described in Bachem et al. (Plant Journal 1996). Tissue samples were taken of stages corresponding to growing stolons and the early stages of tuber formation."

BASE COUNT
ORIGIN 129 a 129 c 158 g 177 t

alignment_scores: 891.50 Length: 191
Ratio: 4.980 Gaps: 1
Percent Similarity: 93.717 Percent Identity: 88.482

alignment_block:
US-09-508-418-2 x BE343078 ..

Align seg 1/1 to: BE343078 from: 1 to: 593

17 SerSerSerPro...LeuAlaPheLeuAsnArgThrSerPheLepr 32
|||||
21 TCCTCCCTCATCGCGTCATTTTAAACGTCACAAATTCATTTC 70
|||||
32 PheSerSerLeuSerArgAsnSerValAsnCySaanglyTTPARGT 49
|||||
71 TTACTTTTCACCTCCCAAGCGCATATGTCATTCATGCTGAGAGA 120
|||||

49 hrArgCysSerValAlaIysAspTyrThrValProSerSerAlaValAsp 65
|||||
121 CAGATGTCCTCGTTGCCAAGATTATACAGTTCCTCCCTCGGAAGTCGAC 170
66 GlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSe 82
|||||
171 GGTAATCAGTTCCTCCGAGCTGATGTGTGTGTGTGTGTGTGTGTGTGT 220
82 rGlyLeuCysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMet 99
|||||
221 TGGACTCTGCTATGCTAAGTGTATTCGGATATATATCCCAATTTGATGG 270
99 alThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArg 115
|||||
271 TGAAGGAGGAGGAGGATGCTGCGGTGGAACAATACGAGCGGTGAAACA 320
116 AspGlyTyrLeuTyrGluGluGlyProAsnSerPheGlnProSerAsp 132
|||||
321 GATGATACCTTATGGAGAGAGGTCTTACAGATTCCACCTTCGGATCC 370
132 oMetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeu 149
|||||
371 TATGTTACATGCTGATGATTTGTGATTTGAAGATTTGTGTGG 420
149 LysAspProAsnAlaProArgPheValLeuTyrPheGlyLysLeuArgPro 165
|||||
421 GAGATCCTGATGCGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 470
166 ValProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePr 182
|||||
471 GTTCCGCGCAAGCTCAGTATCTCCCTTCTTGTGATTTGATGATGATCC 520
182 oGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerPro 199
|||||
521 TGGCAAGCTCAGAGCTGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 570
199 roGlyHisGluGluSerValGlu 206
|||||
571 CAGGTTATGAGCATCAGTTGAG 593
seq_name: gb_est50:AM695211

seq documentation block:
LOCUS AM695211 640 bp mRNA EST 15-JUN-2000
DEFINITION NF092306ST1F1051 Developing stem Medicago truncatula cDNA clone
NF092306ST 5', mRNA sequence.
ACCESSION AM695211
VERSION AM695211.1 GI:7569973
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 640)
He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula stem library
JOURNAL COMMENT
Unpublished (2000)
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 640 Std Error: 0.00
Plate: 092 row: G column: 06
Seq primer: TCACACGAGAAACAGCATATGAC.
FEATURES
Source location/Qualifiers
1..640

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF092306ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"

BASE COUNT 174 a 126 c 146 g 194 t
ORIGIN

alignment_scores:
Quality: 838.00 Length: 214
Ratio: 4.169 Gaps: 1
Percent Similarity: 93.925 Percent Identity: 76.636

alignment block:
US-09-508-418-2 x AM695211 ..

Align seq 1/1 to: AM695211 from: 1 to: 640

300 AlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTyrPheLys 316
|||||
1 GCATTTCTGCTAGTGTAGTAAACAAGTGAAGTATCTTGTGAAGCTTT 50
316 rSerIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrP 333
|||||
51 AAGTATTAGTAAACCTGATAGTGAGAGTACAGTTTGATGATGAACCC 100
333 roGluLysValAlaSerLeuGlnSerArgSerIleValMetThrValPro 349
|||||
101 CAGAGAGAGTGTGTTCTTGCAGAGCAAGACTGTTATCATGACCATTCCT 150
350 SerTyrValAlaSerSerHisLeuArgProLeuSerValAlaAlaAlaAs 366
|||||
151 TCCCAAGTGTGCTAGTCTGCTGACCATCTGTCTCTGCTGTCGAGA 200
366 PalLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerT 383
|||||
201 TGCATTTCAAGTTTATTTATTCCTCCAGTTGCCGACATTTCCATTTCT 250
383 YrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLys 399
|||||
251 ATCCAAAGAGAGCCATTTAGATCAGAAATGCTGATGATGAGGAGTTGAG 300
400 GlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyTh 416
|||||
301 GGAATTGGTCAATTGCAATCCAGCCAGCAAGAGATGCAACATTAGAAC 350
416 rIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyValVal 433
|||||
351 TATATACACCTCATCCCTTTTCCCTAACCGAGCACCACTGAGGGGCTC 400
433 eueLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLys 449
|||||
401 TACTTGTGAATTACATGAGGAGGCTTACCAATTCGGATTTTATCAAG 450
450 ThrGluSerGlnLeuValGluValAlaValAspArgAspLeuArgLysMet 466
|||||
451 ACGAGAGGTGAACCTTTGCAACAGTTGATGACGATTTGAGAAAAAATCT 500
466 uIleLysProLysAlaGlnAspProLeuValValGlyValArgValTrrp 483
|||||
501 TATTAACCAAAATGCTCAGATCCAAATTTTGGGGGTTAGATGTGGC 550
483 roGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThr 499
|||||
551 CTCAGAGCTATTCACAGTTCTTGTGAGACATCTTGATCTTCATGAGCT 600
500 AlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeu 512
|||||
601 GCTAAAGTTTCTTAATAAAATCTGAGATTGAGGGGCTT 640

```

seq_name: gb_est199:BG320666

seq_documentation_block:
LOCUS      BG320666          780 bp      mRNA          EST          27-FEB-2001
DEFINITION  Zm04_06e01_R_Zm04_AAPC_ECORC_cold_stressed_maize_seedlings Zea mays
            CDNA clone Zm04_06e01, mRNA sequence.
ACCESSION  BG320666
KEYWORDS   BG320666.1  GI:13150344
VERSION    EST.
SOURCE     Zea mays.
ORGANISM   Zea mays.
REFERENCE  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
AUTHORS    Singh, J.A., Makul, K., Couroux, P., De Moors, A., Harris, L.J., Hattori,
            J.I., Ouellet, T., Robert, L.S., Sprout, D., and Tinker, N.A.
TITLE      Expressed Sequence Tags from Cold-Stressed Maize Seedlings
JOURNAL    Unpublished (2001)
COMMENT    Contact: Singh, J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-Food Canada
            960 Carling Avenue, Bldg. 20, Ottawa, Ontario, K1A 0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
            Location/Qualifiers
            1..780
            /organism="Zea mays"
            /cultivar="C0328"
            /db_xref="taxon:4577"
            /clone="Zm04_06e01"
            /clone_id="Zm04_AAPC_ECORC_cold_stressed_maize_seedlings"
            /tissue="leaf, crown"
            /note="Vector: Bluescript SK-/XhoI-EcoRI; Site_1: Eco RI;
            Site_2: Xho I; Lower temperature 50 C / hour from 22 to
            120C; bring to 50 in 1 hour from 120C. Leave at 50C 2 days
            , photoperiod 16 hours. Light intensity was 125 uE-1.
            Library prepared by in vivo mass excision from amplified
            library."
BASE COUNT  185 a      198 c      186 g      175 t      36 others
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

201 AGCGCGCTTCATCCGCCGCTCCYCCAGAGCCGCCAAGAGTARTTG 250
206 LUGLNPHEVALARGARGANLEUGLYGLUVALPHEGLUARGLEULE 222
251 AGAGATGCTTCGCCCAACCTGCTGCTGAGCKCTCGAGCSCCTCAAT 300
223 GUUPROPHCYSSERGLYVALTYRVALGLASPPROSERLYSLEUSERME 239
301 KAGCCTTTCTGCTCARGTGTCTATTCTGTGTCCTTCTTARCYCAGCAT 350
239 TLYSALAALPHEGLYLYSVALTYRPLYSLEUGLUGLUTHRGLYSERI 256
351 GAAGGCTCATTTGGAGAGGTTTKCGGTTGGAGAAATCGAGTAGTA 400
256 LELLEGLGLYTHRPHLYSALAILELYSGLUARGSERTHRPROLYS 272
401 TTATTKTGAGAAC. ATCAAGACAAATTCAKAGAGAGCAGACAGATCCAAA 449
273 ALAPROARGAPPROARGLEUPROLYSPROLYSGLYGLINTHRVALISYE 289
450 CCACGADGAGATGCYCGCTTCGAGGCAAGCAAAARGCAGACATTCGATC 499
289 IPHEARGLYSGLYLEUARGMETLEUPROASPALAIISEERLARGLEUG 306
500 TTTGAGGAAGGGTCTTGCCAMGCTTCCAAATTCCAATRATCCAGCTTGG 549
306 IYSELYSLEULYSLEUSERTHRPLYSLEUSERLERTHRYSSEGLU 322
550 GTAGTAAAGTCACAACTATCAGWGAAGACACGACATTCACAAATCAGAT 599
323 LYSGLYGLYTHRISLEUTHRTHYGLUTHRPROGLUGLYVALISERLE 339
600 GACAAAGGATATGTTTGAGTATTAAGCCAAAGGGTGTGTTGGT 649
339 UGLNSETARGSERILEVALMETHRVALPROSERTYRVALALASERASNI 356
650 GCAGCTAAAGTTTATMCATGACTATTCATATGTTGCTAGACACA 699
356 LELUARGPROLEUSERVALALALALALASP. ALALEUSERASPHETRY 372
700 WTTTGGCTCCACTTTCAGACATGCTCAGAGCCCTATC...AAGATCTA 746
372 RTYRPROVALGLYVALAVAL 379
747 TTAAACACCGGTTCTCGCGTA 768

seq_name: gb_est199:BG352959

seq_documentation_block:
LOCUS      BG352959          590 bp      mRNA          EST          01-MAR-2001
DEFINITION  sab93f10.y1 Gm-cl040 glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-cl040-2588.5, similar to SW:POC_ARATH P55826
            PROTOPORPHYRINOGEN OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION  BG352959
KEYWORDS   BG352959.1  GI:13181619
VERSION    EST.
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
            (bases 1 to 590)
AUTHORS    Shoemaker, R., Klein, P., Vodkin, L., Espeland, J., Coryell, V., Khanna,
            A., Balle, B., Marrs, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
            Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers,
            Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk,
            R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann,
            R., Waterston, R., and Wilson, R.
TITLE      Public Soybean EST Project
JOURNAL    Unpublished (1999)
COMMENT    Contact: Shoemaker R./Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine

```

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
Fax: 314 286 1810

Email: estewatson.mustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com High quality sequence stop: 481.

FEATURES

source

```
1 590
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1040-2588"
/clone_lib="Gm-c1040"
/tissue_type="Hypocotyl and Plumule, germinating seeds"
/lab_host="DH10B"
/notes="Vector: pT73Pac (Pharmacia); site_1: EcoRI; site_2: NotI; this cDNA library was constructed from mRNA isolated from hypocotyl and plumule tissues of seeds germinated for three days of the cultivar Williams 82. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a NotI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestions with EcoRI and NotI. The cDNA fragments were directionally cloned into the EcoRI-NotI restriction site of the pT73-Pac vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker."
```

BASE COUNT 156 a 117 c 147 g 170 t
ORIGIN

alignment_scores:

Quality: 797.00 Length: 191
Ratio: 4.478 Gaps: 1
Percent Similarity: 93.194 Percent Identity: 80.105

alignment_block:

US-09-508-418-2 x BG352959 ..

Align seg 1/1 to: BG352959 from: 1 to: 590

```
190 GATGTTGTTGCTGCGAATTCGCTCCTCCATGAGGATCGGTTGA 64
15 GGTGCGCTTGGAAATTCGCTCCTCCATGAGGATCGGTTGA 64
206 uginphevalargarganleuglygluvalphegluarqleuileg 223
65 AGAGTTGTTGCTGCGAATTCGCTCCTCCATGAGGATCGGTTGA 114
223 luepofhecysserglyvaltyrvalglyaspserleuysersermet 239
115 AGCCTTTTGTTCAGGGGCTCTATGCGAGGATCCTCAAAATTAAGTATG 164
240 lysalalaphheglylysvaltrpylesleugluinlthrglyseril 256
165 AAGAGCAGCATTCGGAAGTTGGAGCTGGAAAAAATGCTGTAGCAT 214
256 eilegilygltyrphelysalalleysgluarqserleuysa 273
215 TATTTGGTGGAACTTCAAGACATACAGAGAGAAATGAGCTTCAAAAC 264
273 laproargaspproargleuproluysprolyslyglinthryvalglyser 289
265 CACTCGAGATCCGCGTCCGCAAAACCAAAAGTCAGACTGTGGATCT 314
290 phearglysglyleuargmetleuproaspalalleseralargleucl 306
315 TTCGGGAAGGACTTACCATGTGCTGATGCAATTTCTGCAGACTAGG 364
```

```
306 yserlyseuylseuSertrpylesuSerlethrylsSerGluL 323
365 CACACAAAGTAAAGTTATCTTGGAGCTTCAAGTATTAACAGGANA 414
323 yselglytyrhisleuthrtyrgluthrprogluylvalvalserleu 339
415 GTGAGAGTACAGTTGACATATGAACACAGAAAGAGTGGTTCTTGG 464
340 GlnSerArgSerleValmethrValproserTyValAlaSeranil 356
465 CACTGCAAAACGTGTCTCGACCATTCCTCTATGTTGCTAGTCAT 514
356 eleuargproleuSerValAlaAlaAlaaspAlaSeranpheyTYT 373
515 GCTGCGCTCTGTCTGCTGCTGCTGCTGCAATCTTCAAGTTTAT 564
373 yrPro.ProValGlyAlaVal 379
565 ACCCTTCACAGTGCCTGACGTT 585
```

seq_name: gb_est74:BE455801

seq_documentation_block: 580 bp mRNA

LOCUS BE455801 EST 27-FEB-2001

DEFINITION HVSMEG0015D05f Hordeum vulgare pre-anthesis spike EST library

HVCNDA0008 (white to yellow anther) Hordeum vulgare cDNA clone

HVSMEG0015D05f, mRNA sequence.

ACCESSION BE455801

VERSION BE455801.2 GI:13154672

KEYWORDS EST.

SOURCE barley.

ORGANISM Hordeum vulgare

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

; Triliceae; Hordeum.

1 (bases 1 to 580)

Wing,R., Close,T.J., Kleinofe,A., Wise,R., Begum,D., Fritsch,D., Yu

,Y., Anderson,H., Dale,J., Henry,D., Kennode,S., Palmer,M., Rambo

,T., Saski,C., Schwartzbeck,J., Simons,J., Choi,D.W., Main,D. and

Wood,T.

Development of a genetically and physically anchored EST resource

for barley genomics

Unpublished (2000)

On Jul 26, 2000 this sequence version replaced gi:9465836.

Contact: Wing RA

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Seq primer: AATTAACTCTCACTAAGGCG

High quality sequence stop: 499.

FEATURES

source

```
1 580
/organism="Hordeum vulgare"
/db_xref="taxon:4513"
/clone_lib="HVSMEG0015D05f"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library"
HVCNDA0008 (white to yellow anther)
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"
/notes="Vector: lambda2AP, site_1: EcoRI, site_2: XhoI"
```

BASE COUNT 156 a 124 c 150 g 147 t

ORIGIN

alignment_scores: Quality: 778.00 Length: 192
Ratio: 4.346 Gaps: 0
Percent Similarity: 93.229 Percent Identity: 75.000

alignment block:

US-09-508-418-2 x BE455801 ..

Align seg 1/1 to: BE455801 from: 1 to: 580

```

217 ValPheGluArgLeuIleGluProPheGlySerGlyValIleValGlyAs 233
    |||||
3   GTCTTTGAGCGCTTATGAACTTTCTGCTCAGGTGTGTGCTGGTGA 52
233 pProSerLysLeuSerMetLysAlaAlaPheGlyLysValIlePheLys 250
    |||||
53 TCCTTCCAGGCTCAATATGAAAGGCTGCTTGGGAGAGTTGGAGGTTGG 102
250 IuGluThrGlyGlySerIleIleGlyIleThrPheLysAlaIleLysGlu 266
    |||||
103 AGGAGATTGGGGGTAGTATTTGTTGGTGAACATCAAGGCATCAGGAT 152
267 ArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 283
    |||||
153 AAAGGGAAGAACCCCAACCCCAAGGATCCCGACTCCGACATCCGACCAA 202
283 sGlyGluThrValGlySerPheArgLysGlyLeuArgMetLeuProAsp 300
    |||||
203 GGGACAGACGGTGGCATCTTCAAGAGGCTCTGGCCATGCTCCGGAATG 252
300 IalleSerAlaArgLeuGlySerLysLeuLysLeuSerThrPheLys 316
    |||||
253 CCATGCGCTCTAGGTGGGTAGTAAAGTCAAGCTGTCATGGAACTTACG 302
317 SerIleThrLysSerGlyLysGlyIleValIleThrLysLeuThrPro 333
    |||||
303 AGCATTCACAAAGCGGACACCAAGCATATGTTAGCTTATGAAACACC 352
333 oGluGlyValIleSerLeuGlnSerArgSerIleValMetThrValPro 350
    |||||
353 AGAAGGACTGTTTCAGTGCAGCTAAAGTATCATCATGACCATCCCGT 402
350 eTyTyValAlaSerAsnIleLeuArgProLeuSerValAlaIleAlaAsp 366
    |||||
403 CATATGTTGCTAGTATCTTACGCCACTTCAATTGATGACGACGAT 452
367 AlaleuSerAsnPheTyTyProProValGlyValAlaValThrIleSer 383
    |||||
453 GCACCTCTCAAAATCTATATCCGCCAGTGTCTCTAATCTGTTTATA 502
383 rProGluGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLys 400
    |||||
503 TCCAAAGAAAGCTATTAGAAAGAACTCTTATGATGGGAAAGCTCAGG 552
400 lYPheGlyGlnIleuHisProArgThr 408
    |||||
553 GGTTCGGCGACGTGCATCCACGTAGC 578
seq_name: gb_est39:AV538014
seq_documentation block:
LOCUS AV538014 596 bp mRNA EST 06-SEP-2000
DEFINITION AV538014 Arabidopsis thaliana roots Columbia Arabidopsis thaliana
CDNA clone RZ109a12F 3', mRNA sequence.
ACCESSION AV538014
VERSION AV538014.1 GI:8698297
KEYWORDS EST.
SOURCE
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 596)
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL

```

MEDLINE 20363093

COMMENT

Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES

source

1..596

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="RZ109a12F"

/clone_id="Arabidopsis thaliana roots Columbia"

/tissue_type="roots"

/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:

BASE COUNT

165 a

135 c

126 g

170 t

ORIGIN

alignment_scores:

Quality: 739.00 Length: 177
Ratio: 4.452 Gaps: 0
Percent Similarity: 93.785 Percent Identity: 79.096

alignment_block:

US-09-508-418-2 x AV538014/rev ..

Align seg 1/1 to reverse of: AV538014 from: 1 to: 596

```

372 TyTyProProValGlyValAlaValThrIleSerTyTyProGluAlaI 388
    |||||
591 TATTACCCACAGTTCAGACAGATCTATCTGATCCGACCAAGAAAGCAT 542
388 eArgAspGluArgLeuValAspGlyGluLeuLysGlyPheGlyGlnIle 405
    |||||
541 CCGAACAGAAATGTTGATGATGCTGAACCTAAAGGTTTGGCGCAATTGC 492
405 lSProArgThrGlnGlyValGluThrLeuGlyThrIleTySerSer 421
    |||||
491 ATCCACGACGACGAGAGGTTGAACATTTAGAACTATCTACAGCTCTCA 442
422 LeuPheProAsnArgAlaProLysGlyArgValLeuLeuAsnTyTl 438
    |||||
441 CTCTTCCAAATCGCCACGCCCGGAAAGAAATTTGCTGTGAACATCAT 392
438 eGlyGlyAlaLysAsnProGluIleLeuSerLysThrGlnSerGlnLeu 455
    |||||
391 TGGCGGCTCTACAAACCGGAATTTCTGCCAAGTCTGAAGGTGAGTTAG 342
455 aGluValAlaAspArgAspLeuArgLysMetLeuIleLysProLysAla 471
    |||||
341 TGGAGACAGATTGACAGAGATTGAGGAAATCTAATTATAGCTTAATTCG 292
472 GluAspProLeuValValGlyValArgValIleProGlnAlaIlePro 488
    |||||
291 ACCGATCCACTTAATTTAGAGTTAGGATGGGATGCCCTCAAGCATCTCA 242
488 nPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysAlaAlaMet 505
    |||||
241 GTTTCAGTGGTCACTTGTATATCTTGACAGCGCTAAATCATCTCTA 192
505 sAspAsnGlyLeuGlyLeuPheLeuGlyLysTyValSerGly 521
    |||||
191 GCTCTTGGGCTACGAGAGGCTATTTTGGGGGCAATTAGCTGCTGGT 142
522 ValAlaLeuGlyArgCysValGluGlyAlaTyArgIleValAlaSerGlu 538
    |||||
141 GTAGCCTTAGCGCGGTGTGTAGAAAGCGCATATGAAACCGGATTGAGGT 92
538 lThrGlyPheLeuSerArgTyAlaTyLys 548
    |||||
91 CAACAACCTCATGTCAACGCTTACAG 61

```

```

seq_name: gb_est190:BF624024
seq_documentation_block:
LOCUS      BF624024      510 bp      mRNA      EST      21-FEB-2001
DEFINITION HVSMA00081L6f Hordeum vulgare seedling shoot EST library
            HVCDA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMA00081L6f,
            mRNA sequence.
ACCESSION  BF624024
VERSION    BF624024.2  GI:13082749
KEYWORDS   EST.
SOURCE     barley.
ORGANISM   Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
            ; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 510)
            Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
            , Y., Anderson, H., Dale, J., Henry, D., Kerndle, S., Palmer, M., Rambo
            , T., Saski, C., Schwartzbeck, J., Simmons, T., Choi, D.W., Main, D. and
            Wood, T.
            Development of a genetically and physically anchored EST resource
            for barley genomics
            Unpublished (2000)
            On Dec 18, 2000 this sequence version replaced gi:11887758.
TITLE      Development of a genetically and physically anchored EST resource
            for barley genomics
JOURNAL    Unpublished (2000)
COMMENT    Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: AATTACCTCTACTAAAGG
            High quality sequence stop: 505.
            Location/Qualifiers
                1..510
                /organism="Hordeum vulgare"
                /cultivar="Morex"
                /db_xref="taxon:4513"
                /clone="HVSMA00081L6f"
                /clone_1lb="Hordeum vulgare seedling shoot EST library
                HVCDA0001 (Cold stress)"
                /tissue-type="Seedling shoot"
                /lab_host="TJG121"
                /note="Vector: lambdaZAP; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      148 a      110 c      116 g      135 t      1 others
ORIGIN
Alignment_scores:
    Quality:      685.00      Length:      168
    Ratio:        4.363      Gaps:        0
    Percent Similarity: 93.452      Percent Identity: 77.381
alignment_block:
US-09-508-418-2 x BF624024 ..
Align seg 1/1 to: BF624024 from: 1 to: 510
320 LysSerGluLysGlyLeuTyrHisLeuThrTyrGluThrProGluGlyVala 336
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
6 AAGCGGACACCAAGATATGTAATGATTGTAACACCAAGAGGACT 55
336 lvalserleuglnserargserillevalmethyrvalproserlyrvala 353
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
56 TGTTCAGTCAGCGGCTAAAGTGTATCATCCGATCCGTCATATGTTG 105
353 lserasnlleuuarqproleuservalalalalalalaspalaleuser 369
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
106 CTAGGATATCTTACGCCACTTTCAATGATGACGACGATCATCTCA 155
370 Asnphetyrtyr-proprovalalalavalthrilesertyrproglngl 386
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

156 AATCTATTATCCGCCAGTGTGCTGTAAGTTCATATCCAAAGA 205
386 ualalleargaspgluarleuvalaspolygluleuylslypnehyg 403
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
206 AGCTATTAGAAAAGAAATGCTTATGATGGGAGCTCCAGGGGTTTCGGCC 255
403 lneuhsiproargthrngllyvalgluThrleuglythriletyrser 419
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
256 AGCTGCATCCACGTACCCAAAGAGGTGACACTTACGACATATATAC 305
420 SerSerleupheproasnargalaprolysglyargvalleuleuans 436
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
306 TCTTCCTCTCTTCCCAATCGTCTGCTCGAAGAGTGTACTTCTGAA 355
436 nryrllleglylvalalysasnproglulleuserlystrnglsersg 453
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
356 CTATATCGGGGGTTCACAAATACAGGAGATCGTCCAAAGCCGAGAGTG 405
453 lneuvalgluvalalaspargaspheuarqlysmetleuiletypro 469
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
406 ACTTAGTAGAAGCTGTGTGATCGTGCATCTCAGAAATGTTGATTAACCT 455
470 lysalaglaspheproleuvalalgllyvalargvaltrproglalal 486
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
456 AGAGCAGACAGACCTTAGCATTAGGGGTACAGTGTGCCACAAACCAAT 505
486 epro 487
|||||
506 ACCA 509
seq_name: gb_est42:AM119581
seq_documentation_block: 498 bp      mRNA      EST      17-JUL-2000
LOCUS      AM119581
DEFINITION sda46d01.y1 gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            GM-cl016-3002 5' similar to SW:PF0C_ARATH P53826 PROTOPORPHIRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION  AM119581
VERSION    AM119581.1  GI:6094914
KEYWORDS   EST.
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
            1 (bases 1 to 498)
REFERENCE  1
            Shoemaker, R., Keim, P., Vodkin, L., Erpelund, J., Corvill, V., Khanna
            , A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
            Wylie, T., Underwood, K., Steptoe, M., Rheising, B., Allen, M., Bowers
            , Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schmitz
            , R., Riltter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann
            , R., Waterston, R. and Wilson, R.
            Public Soybean EST Project
            Unpublished (1999)
TITLE      Public Soybean EST Project
JOURNAL    Unpublished (1999)
COMMENT    Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1114 Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 425.
            Location/Qualifiers
                1..498
                /organism="Glycine max"
                /db_xref="taxon:3847"
FEATURES
SOURCE

```

```

/c1one="GENOME SYSTEMS CLONE ID: Gm-c1016-3002"
/c1one_lib="Gm-c1016"
/tissue_type="Immature flowers of field grown plants"
/lab_host="X110-Gold"
/note="Vector: pBluescript II XR, Site_1: EcoRI, Site_2:
XhoI. This cDNA library was constructed from mRNA isolated
from Immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into X110-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelding."
BASE COUNT      137 a      106 c      111 g      143 t      1 others
ORIGIN

```

```

alignment_scores:
  Quality: 674.00      Length: 164
  Ratio: 4.266      Gaps: 0
  Percent Similarity: 96.341      Percent Identity: 82.317

```

```

alignment_block:
  US-09-508-418-2 x AW119581 ..

```

```

Align seg 1/1 to: AW119581 from: 1 to: 498

```

```

328 LeuTrrTYrGluThrProGluGlyValValSerLeuGlnSerArgSerI 344
      |||||||
8 TTGACATATGAAACACCCAGAGGAGTGTCTTTCGACCAAAAGCTGT 57
344 eValMetThrValProSerTYrValAlaSerAsnIleLeuArgProLeuS 361
      |||||||
58 TGCTCTGACCATCTCTTCTTCTTATGTTGCTAGTACATGCTGCTCTGT 107
361 eValAlaAlaAlaAlaAspAlaLeuSerAsnPhetyTYrProProValGly 377
      |||||||
108 CTGCGTGTGCTGACAGATGCACTTCAAAAGTTTATTTACCCCGCATGCT 157
378 AlaValThrIleSerTYrProGlnGluAlaIleArgAspGluVal 394
      |||||||
158 GCAGTTTCCATATCTCTATCCAAAGAGCTATTAGACAGAAATGCTTGAT 207
394 IAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGly 411
      |||||||
208 AGATGCTAGTTGAAGGGGTTGGTCAATTGCAATGCCAGTACGCAAGAG 257
411 aGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAla 427
258 TGGAAACATTTAGGAACATATATACAGCTCATCTATTTCCCAACCGACA 307
428 ProLysGlyArgValLeuLeuAsnTYrIleGlyAlaLysAsnPr 444
308 CCACCTGGAAGGGTTCATCTTGAATTACATTTGAGAGAGCAACTAATAC 357
444 oGluIleLeuSerLysThrGlnSerGlnLeuValGluValAlaAspArg 461
358 TGGAAATTTTATCGAAGACGACAGTGAACCT.GTGGAAACAGTTGATGAG 406
461 sPLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVal 477
407 ATTGAGAGAAATCTTATTAACCAATGCGCCAGATTCATTGTGAGTIN 456
478 GlyValArgValTTPProGlnAlaIleProGlnPheLeuVal 491
457 GGGGTGAGACTGTGGCTCAAGCTATTTCACAGTTTCTTACTT 498

```

```

seq_name: gb_est69:BE021686

```

```

seq.documentation_block:
LOCUS      BE021686      480 bp      mRNA      EST      21-NOV-2000
DEFINITION sm61b05.y1 Gm-c1028 glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1028-8362 5' similar to SW:PPOC.ARAH P55826 PROTOPORPHYRINOGEN
            OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION  BE021686
VERSION    BE021686
KEYWORDS   BE021686.1 GI:8284127
SOURCE     soybean.
ORGANISM   Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
            1 (bases 1 to 480)
REFERENCE  Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Corryell,V., Khanna
            A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            R., Waterson,R. and Wilson,R.
            Public Soybean EST Project
            Unpublished (1999)
            Contact: Shoemaker R./Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: este@wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Insert Length: 1299 Std Error: 0.00
            High quality sequence stop: 394.
            Location/Qualifiers
            1..480
            /organism="Glycine max"
            /db_xref="taxon:3847"
            /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-8362"
            /clone_lib="Gm-c1028"
            /tissue_type="roots of 'Supernod' plants"
            /lab_host="DH10B"
            /note="Vector: pBluescript II XR, Site_1: EcoRI, Site_2:
            XhoI. The mRNA was isolated from roots of Glycine max
            'Supernod' plants generously donated by Dr. Gary Stacey.
            The seedlings were inoculated with Bradyrhizobium
            japonicus, strain USDA110 prior to harvest. Stratagene's
            cDNA synthesis kit (catalog number 200401) was used to
            synthesize the cDNA. First-strand synthesis was performed
            with 5-methyl dCTP, hence the ligated cDNA was
            hemimethylated. A modification of Stratagene's
            first-strand synthesis primer was used. An 'anchor'
            nucleotide (V-A, C, or G) was added to the 3' end of the
            primer [GAGACGAGAGAGAGAGAGACTGTCGAG(T)18V] to anchor
            the primer at the 5' end of the poly(A) tract. After
            second-strand synthesis, the cDNA ends were filled in with
            cloned Pfu DNA polymerase, ligated to EcoRI adapters and
            subsequently phosphorylated. The XhoI site within the
            first-strand synthesis primer was then restricted by
            digestion with XhoI; all XhoI sites in the cDNA would be
            protected by their hemimethylated status. The cDNA
            constructs were size-fractionated with a 500bp cutoff,
            using G100BHR life Technologies' cDNA Size Fractionation
            column. The column eluent was then ligated into
            Stratagene's pBluescript II XR predigested vector
            (pBluescript II SK(+)) that has been digested with EcoRI
            and XhoI, and phosphorylated by Stratagene). Both the
            white and blue colonies appear to contain recombinant
            plasmids with cDNA inserts, based on size (n=25). This
            library was constructed by Dr. Paul Kelm and Dr. Virginia

```

BASE COUNT 138 a 98 c 104 g 137 t 3 others
ORIGIN

alignment_scores:

Quality: 629.00 Length: 158
Ratio: 4.338 Gaps: 0
Percent Similarity: 91.772 Percent Identity: 77.848

alignment_block:

US-09-508-418-2 x BE021686 ..

Align seg 1/1 to: BE021686 from: 1 to: 480

```

297 LeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTr 313
|||||.....|.....|.....|.....|.....|.....|.....|.....|
7 TTGGCTGATGCAATTTTCCAGACTAGCCAAACAAAGTAATATCTTG 56
313 pLysLeuSerSerIleThrLysSerGluLysGlyLysThrHisLeuThr 330
|||||.....|.....|.....|.....|.....|.....|.....|.....|
57 GAAGCTTCAAGTATTAGTAAACGAGTAGGAGAGAGAGAGAGAGAGAGAT 106
330 YRGLuThrProGluGlyValValSerLeuGlnSerArgSerIleValMet 346
|||||.....|.....|.....|.....|.....|.....|.....|.....|
107 ATGAAACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 156
347 ThrValProSerTrpValAlaSerAsnIleLeuArgProLeuSerValAl 363
|||||.....|.....|.....|.....|.....|.....|.....|.....|
157 ACCATTCCTCTCTATGCTAGTACATGCTGCTGCTGCTGCTGCTGCTG 206
363 AlaAlaAspAlaLeuSerAsnProLeuThrTrpProValGluValAlaVal 380
|||||.....|.....|.....|.....|.....|.....|.....|.....|
207 TGCTGCAATGACCTTCAAGATTATATACCTCCAGAGAGAGAGAGAGAGAT 256
380 hrLysSerTrpProGluGlyValAlaLeuArgAspGluArgLeuValAspGly 396
|||||.....|.....|.....|.....|.....|.....|.....|.....|
257 CCATATCCTATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 306
397 GluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThr 413
|||||.....|.....|.....|.....|.....|.....|.....|.....|
307 GAGTTGAAGGGGTGGTCAATTCGATCCAGTACGCAAGAGAGAGAGAGAGAT 356
413 rLeuGlyThrIleTrpSerSerSerLeuPheProAsnArgAlaProLysG 430
|||||.....|.....|.....|.....|.....|.....|.....|.....|
357 ATTANGAACTATATACAGCTCATCTATTCGCCAACCGAGACACCACTG 406
430 IYArgValLeuLeuLeuAsnTrpIleGlyAlaAlaLysAsnProGluThr 446
|||||.....|.....|.....|.....|.....|.....|.....|.....|
407 AAGGGTCTACTCTNGAATTACATTGAGAGAGAGAGAGAGAGAGAGAGAT 456
447 LeuSerLysThrGluSerGlnLeu 454
|||||.....|.....|.....|.....|.....|.....|.....|.....|
457 NTATCAAGACGACAGTGAACCTT 480

```

seq_name: gb_est39:BG355389

seq_documentation_block:

LOCUS BG355389 479 bp mRNA EST 06-MAR-2001
DEFINITION 947043H05.y2 947 - 2 week shoot from Barkan lab Zea mays CDNA, mRNA
sequence.

ACCESSION BG355389 GI:13237375

KEYWORDS EST.

SOURCE Zea mays.

ORGANISM Zea mays

REFERENCE 1 (bases 1 to 479)

AUTHORS Walbot V.

TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University

JOURNAL:

Unpublished (1999)

COMMENT:

Contact: Walbot V

Department of Biological Sciences

Stanford University

855 California Ave, Palo Alto, CA 94304, USA

Tel: 650 723 2227

Fax: 650 725 8221

Email: walbot@stanford.edu

Plate: 947043 row: H column: 05.

Location/Qualifiers

1. 479

/organism="Zea mays"

/cultivar="B73"

/db_xref="taxon:4577"

/clone_lib="947 - 2 week shoot from Barkan lab"

/rissue_type="leaf and stem, including leaf base"

/dev_stage="2 week Old seedling (3 leaves)"

/lab_host="X11-Blue"

/note="Organ: shoot; Vector: Lambda ZAP (pBluescript SK-);

Site 1: EcoRI; Site 2: XhoI; Directionally cloned using

Stratagene's UniZap XR cDNA cloning kit with the 5' end

at the EcoRI site. The library represents 8 x 10e5

independent recombinant phage. The plants were greenhouse

grown."

BASE COUNT 122 a 108 c 126 g 123 t

ORIGIN

alignment_scores:

Quality: 621.00 Length: 151
Ratio: 4.500 Gaps: 0
Percent Similarity: 91.391 Percent Identity: 79.470

alignment_block:

US-09-508-418-2 x BG355389 ..

Align seg 1/1 to: BG355389 from: 1 to: 479

```

398 LeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLe 414
|||||.....|.....|.....|.....|.....|.....|.....|.....|
2 CTCACAGGCTTTGGCCAGTGGCATCCAGCAGTACAGAGAGAGAGAGAT 51
414 uGlyThrIleTrpSerSerSerLeuPheProAsnArgAlaProLysGly 431
|||||.....|.....|.....|.....|.....|.....|.....|.....|
52 AGGAACAATATACAGTCCCTCACTCTTCCAAATCGCTCCGACGGTA 101
431 rGlyValLeuLeuAsnTrpIleGlyAlaAlaLysAsnProGluThrLeu 447
|||||.....|.....|.....|.....|.....|.....|.....|.....|
102 GGGGTACTTCTTAACCTACATAGAGAGGCTCTCAACACACAGAAATTGTT 151
448 SerLysThrGluSerGlnLeuValGluValAlaAspArgAspLeuArgLys 464
|||||.....|.....|.....|.....|.....|.....|.....|.....|
152 TCCAGAGCTGAAGTGAGCTGCTGATATGAGTTCGACCGTACCTCGGAA 201
464 SmetLeuIleLysProLysAlaGlnAspProLeuValAlaGlyValArgY 481
|||||.....|.....|.....|.....|.....|.....|.....|.....|
202 AATGCTTATTAATTTACACAGCTGACCTTATGCTTGGGTGCTCGAG 251
481 alTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeu 497
|||||.....|.....|.....|.....|.....|.....|.....|.....|
252 TTGGCCACAAGCCATACCTCAGCTCTCTGTTAGACATCTTGATCTTCTG 301
498 SerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGlyLeuPheLe 514
|||||.....|.....|.....|.....|.....|.....|.....|.....|
302 GAAGCCCGCAAAAGCTGCCCTGAGCCGAGGCTGCTAGACGTGGCTTCCCT 351
514 uGlyGlyAsnTrpValSerGlyValAlaLeuGlyArgGlyValGluGly 531
|||||.....|.....|.....|.....|.....|.....|.....|.....|
352 AGGAGGAGAACTATGTTGACAGAGAGTTCCTGGCGAGATGCTTGAGGCG 401
531 IaTrpGluValAlaSerGluValThrGlyPheLeuSerArgTrpAlaTrp 547
|||||.....|.....|.....|.....|.....|.....|.....|.....|
402 CGTATGAAGTGGCTCGCAATATCTGACTTCTTGACCAAGTATGGCTAC 451

```


124 roAnSerPhglnProSerAspProMetLeuThr 135
|||||
405 CTAACAGTTTCACGCTTCGATCCTATGTTGACT 439

seq_name: gb_est73:BE326089

seq_documentation block:

LOCUS BE326089 590 bp mRNA EST 21-DEC-2000
DEFINITION NF084G10ST1F1081 Developing stem Medicago truncatula cDNA clone
ACCESSION BE326089
VERSION BE326089.2 GI:11936674

KEYWORDS
SOURCE barrel medic.
ORGANISM Medicago truncatula

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 590)
He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
,C.J., Flores,H.R., Imman,J.T., Weller,J.W., May,G.D. and Dixon
,R.A.

TITLE
JOURNAL Expressed sequence tags from the Samuel Roberts Noble Foundation
COMMENT Unpublished (2000)
On Jul 14, 2000 this sequence version replaced gi:9199866.
Contact: Dixon RA
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380

FEATURES
source
1..590
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF084G10ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"
Seq primer: TCACACGAGAACACGCTATGAC.
Location/Qualifiers

BASE COUNT 130 a 153 c 131 g 176 t
ORIGIN
alignment_scores:
Quality: 552.50 Length: 198
Ratio: 3.635 Gaps: 7
Percent Similarity: 76.768 Percent Identity: 59.596

alignment block:

US-09-508-418-2 x BE326089 ..

Align seg 1/1 to: BE326089 from: 1 to: 590

8 AsnHisProSniLePheThrHisGlnSerSerSerProLeuAlaPh 24
:|||||
21 TCTCCACCT...CTTCCACCCACCTCAACCCCTTCCGCCCTTA..... 62
24 eLeuAsnArgThrSer.PheLeProPheSerSer..... 35
63TCTCCATTCCTCCCTTCCGCCCCCAATTCAAAATCA 99
36TLeSerIysArgAsnSerValAsnGlyTrpArgThrAr 50
100 AAATTCATTGCAATCGCAATCAATCAATAC...CCCATTCCTCGCTCCG 146

seq_name: gb_est51:AW755783

seq_documentation block:

LOCUS AW755783 451 bp mRNA EST 21-NOV-2000
DEFINITION S109c11.y1 Gm-cl036 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl036-813 5', similar to SW:PPCC.ARAH P5826 PROTOPHYRINOMEN
OXIDASE, CHLOROPLAST PRECURSOR ;, mRNA sequence.
ACCESSION AW755783
VERSION AW755783.1 GI:7685135
KEYWORDS EST.
SOURCE soybean.

ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
1 (bases 1 to 451)
Shoemaker,R., Keim,P., Vodkin,L., Erpedding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Stepcoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 894 Std Error: 0.00

REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT

50 gCySerValAlaLysAspTyrThrVal.....ProSerSera 63
|||||
147 TTGCTCATTCGCGAGAGAAATCCACCCGATCTCGACTAAACACCTCAA 196
63 lavalaspGlyGlyProAlaAlaGluLeuaspCysValIleValAla 79
:|||||
197 AATCAAAATCCGGAGAGAGATCTCTCATCCGTTGATGTAGTGTGGTGGT 246
80 GlyIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||||
247 GGCATCAGTGGCTTTGGCATGCTCAGGCTCTTCTTACAAAGCATGCTCA 296
95ProAsnLeuMetValThrGluAlaArgaspArgAlaGlyAla 109
|||||
297 TGGTGTCGCCAATGTTGTTGTTACAGAGCGAAGACCGAGTTGGTGGGA 346
109 snllehrThrValGluArgaspGlyTyrLeuTrpIuGluGlyProAsn 125
|||||
347 ATATTATTACTGTGGAGAGAGATGGGTTCTTGGGAGAGAGTCTTAAT 396
126 SerPhglnProSerAspProMetLeuThrMetAlaValaspCysGlyLe 142
|||||
397 AGTTTCAACCTTCTGATCTATGCTCACCATGCTGTGGACAGTGGTTT 446
142 ulYaspaspLeuValleuGlyaspProAsnAlaProArgPheValleuT 159
|||||
447 GAAGACAGATCTGTTTGGTGCATCTGATCACCCTAATTTGTTGT 496
159 rPlYsolYlyLeuArgProValProSerIysLeuThrAspleuProPhe 175
|||||
497 GGAATGGGAAATGAGCGCGGCGGAAACCAACTGATTTGCCCTTTC 546
176 PheAspleuMetSerIleProGlyIysLeuArgAlaGlyPhe 189
|||||
547 TTTGACTTTAATGATATTGCTGGCAAGATCAAGCTGGCTTT 588

FEATURES High quality sequence stop: 298.
Location/Qualifiers
source 1..451
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1036-813"
/clone_lib="Gm-c1036"
/tissue_type="somatic embryos cultured on MSD 20"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; This
cDNA library was constructed from mRNA isolated from
somatic embryos (age ranging from 2 months to 9 months)
cultured on MSD 20. The library was prepared using the
Life Technologies superscript cDNA library construction
kit. Complementary DNA was synthesized from mRNA using a
poly (dT) sequence with a NotI restriction site. SalI
linkers adapters were ligated to the blunt-ended cDNA
fragments followed by NotI digestion. The cDNA fragments
were directionally cloned into the NotI-SalI restriction
site of the pSPORT1 vector. The ligated cDNA fragments
were transformed into E.coli Electromax DH10B host cells.
This library was constructed in the laboratory of Dr. Lila
Vodkin by Anu Khanna at the University of Illinois at
Urbana-Champaign. e-mail: l-vodkin@uiuc.edu"

BASE COUNT 125 a 88 c 111 g 127 t
ORIGIN

alignment_scores:
Quality: 543.00 Length: 146
Ratio: 4.022 Gaps: 0
Percent Similarity: 92.466 Percent Identity: 75.342

alignment_block:
US-09-508-418-2 x AW755783 ..

Align seg 1/1 to: AW755783 from: 1 to: 451

```
382 SerTYrProGInGluAlaIleArgSPGLuArgLeuValAspGlyGluLe 398
|||||
11 TCCTGTCCAAAGAGCTATTAGATCAGAAATGCTGATAGATGGTGAATT 60
398 uLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuG 415
|||||
61 GAAGGGGTTTGGTCAATTCATCCACGTAGCAAGAGAGCTGGAACATTAG 110
415 lYThrIleTySerSerSerLeuPheProAsnArgAlaProLysGlyArg 431
|||||
111 GAACATATATACAGCTCATCATCTATCCCAACCGACACCTCGGAAGG 160
432 ValLeuLeuLeuAsnTyrlleGlyValAlaLysAsnProGluIleLeuSe 448
|||||
161 GTTCTACTCTTGAATTAATTCAGAGAGCACTAAACTGCAATTTTATC 210
448 rTySThrGlnSerGlnLeuValGluValAlaAspArgAspLeuArgLysM 465
|||||
211 AAAGAGGAGCACTGAACCTGTGAAACAGCTGACATTCGACATTGAGCAA 260
465 eLeuIleLysProLysAlaGlnAspProLeuValAlaGlyValArgVal 481
|||||
261 TCCTTATTAACCCCAATATGCCAGATTCATTGTGAGGGGTGAGACTG 310
482 TrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSe 498
|||||
311 TGGCCCTCAAGCTATTCCACAGCTCTTATTGGCCATATTGATCTTCTAGA 360
498 rThraAlaLysAlaAlaMetAsnAspAsnGlyLeuGlnGlyLeuPheLeuG 515
|||||
361 TGTGGCTAAAGCTCTCATCAGAAATATACGGGTTGAAGGGCTCTTTCTTT 409
515 lYGIyAsnTyValSerGlyValAlaLeuGlyArgCys 527
|||||
410 G.GGGACTTATGTGTCTGGCGGCTGCTTATGACATGC 447
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:33:02 ; Search time 100.76 Seconds

(without alignments)
719.563 Million cell updates/sec

Title: US-09-508-418-2

Sequence: 1 MTTPTIANHPNFTHOSSSS.....EGAYVASEVTGLSRVAYK 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP unclassified:*
13: SP vertebrate:*
14: SP virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-----------|--------------------|
| 1 | 2832 | 99.6 | 548 | 10 Q9SB14 | Q9sb14 nicotiana t |
| 2 | 2824 | 99.3 | 548 | 10 Q9SAY9 | Q9sax9 nicotiana t |
| 3 | 2565.5 | 90.2 | 557 | 10 Q64384 | Q64384 solanum tub |
| 4 | 2061 | 72.5 | 545 | 10 Q9ZS15 | Q9zsl5 arabidopsis |
| 5 | 2004.5 | 70.5 | 555 | 10 Q9SP16 | Q9sp16 cichorium l |
| 6 | 1989.5 | 70.0 | 535 | 10 Q9M629 | Q9m629 zea mays (m |
| 7 | 1966.5 | 69.1 | 562 | 10 Q9LRI8 | Q9lri8 spinacia ol |
| 8 | 1530.5 | 53.8 | 563 | 10 Q9ZTA7 | Q9zta7 chlamydomon |
| 9 | 1527.5 | 53.7 | 563 | 10 Q9T0P9 | Q9t0p9 chlamydomon |
| 10 | 522.5 | 18.4 | 502 | 10 Q9SLW5 | Q9slw5 glycine max |
| 11 | 515 | 18.1 | 436 | 2 Q67814 | Q67814 aquilex aeo |
| 12 | 459 | 16.1 | 544 | 10 Q9FYV8 | Q9fyv8 zea mays (m |
| 13 | 456 | 16.0 | 501 | 10 Q9LYA7 | Q9lya7 arabidopsis |
| 14 | 456 | 16.0 | 547 | 10 Q9FMS9 | Q9fms9 arabidopsis |
| 15 | 424 | 14.9 | 404 | 10 Q64385 | Q64385 solanum tub |
| 16 | 407.5 | 14.3 | 462 | 2 Q9KDX8 | Q9kdx8 bacillus ha |
| 17 | 363.5 | 12.8 | 462 | 2 Q9RV99 | Q9rv99 deinococcus |
| 18 | 317 | 11.1 | 468 | 2 Q69829 | Q69829 streptomyce |
| 19 | 290.5 | 10.2 | 391 | 10 Q9FW53 | Q9fw53 oryza sativ |

| | | | | | |
|----|-------|-----|-----|-----------|--------------------|
| 20 | 254 | 8.9 | 435 | 2 Q9JS40 | Q9js40 chlamydia p |
| 21 | 250 | 8.8 | 475 | 5 Q9UBD0 | Q9ubd0 drosophila |
| 22 | 249 | 8.8 | 435 | 2 Q9Z718 | Q9z718 chlamydia p |
| 23 | 247 | 8.7 | 475 | 5 Q9VC52 | Q9vc52 drosophila |
| 24 | 236.5 | 8.3 | 63 | 10 Q9M5P0 | Q9m5p0 nicotiana t |
| 25 | 231.5 | 8.1 | 424 | 2 Q9PLH9 | Q9plh9 chlamydia m |
| 26 | 211 | 7.4 | 424 | 2 Q84750 | Q84750 chlamydia t |
| 27 | 174.5 | 6.1 | 421 | 2 Q9RSV7 | Q9rsv7 deinococcus |
| 28 | 169.5 | 6.0 | 441 | 2 Q9RTN3 | Q9rtn3 deinococcus |
| 29 | 162 | 5.7 | 524 | 13 Q918A7 | Q918a7 scomber jap |
| 30 | 153.5 | 5.4 | 416 | 2 Q9RI54 | Q9ri54 streptomyce |
| 31 | 148 | 5.2 | 440 | 1 Q58162 | Q58162 pyrococcus |
| 32 | 146 | 5.1 | 430 | 1 Q9HRU3 | Q9hru3 halobacteri |
| 33 | 144 | 5.1 | 488 | 10 Q9LYT1 | Q9lyt1 arabidopsis |
| 34 | 143 | 5.0 | 490 | 10 Q9SKX5 | Q9skx5 arabidopsis |
| 35 | 136.5 | 4.8 | 467 | 2 Q9S2P5 | Q9s2p5 streptomyce |
| 36 | 130 | 4.6 | 890 | 5 Q9VW97 | Q9vw97 drosophila |
| 37 | 127 | 4.5 | 574 | 10 Q9FVR9 | Q9fvr9 arabidopsis |
| 38 | 126 | 4.4 | 489 | 2 Q9KHE1 | Q9khe1 streptomyce |
| 39 | 125 | 4.4 | 499 | 2 Q9K566 | Q9k566 mycobacteri |
| 40 | 125 | 4.3 | 578 | 10 Q9FV57 | Q9fv57 arabidopsis |
| 41 | 121.5 | 4.3 | 443 | 2 Q82865 | Q82865 arabidopsis |
| 42 | 121.5 | 4.3 | 479 | 2 Q9R6X5 | Q9r6x5 anabaena sp |
| 43 | 121 | 4.3 | 404 | 1 Q28536 | Q28536 archaeoglob |
| 44 | 119.5 | 4.2 | 472 | 2 Q9ZBG4 | Q9zbg4 streptomyce |
| 45 | 119.5 | 4.2 | 570 | 10 Q65813 | Q65813 haematococc |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY: | PRT: | 548 AA. |
|--------|--|--------------|------|---------|
| Q9SB14 | Q9SB14 | | | |
| AC | Q9SB14: | | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) | | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update) | | | |
| DE | PROTOPORPHYRINOGEN OXIDASE PX-1 (EC 1.3.3.4). | | | |
| OS | Nicotiana tabacum (Common tobacco). | | | |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; | | | |
| OC | Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; | | | |
| OC | Solanales; Solanaceae; Nicotiana. | | | |
| OX | NCBI_TaxID=4097; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=CV, STR1: | | | |
| RA | Horikoshi M., Mametsuka K., Hirooka T.; | | | |
| RT | "The Molecular Basis of Photobleaching Herbicide Resistance in | | | |
| RT | Tobacco." | | | |
| RT | Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases. | | | |
| DR | EMBL; AF044128; AAD02290.1; - | | | |
| KW | Oxidoreductase. | | | |
| SQ | SEQUENCE 548 AA: 59241 MW: 00E1A106F299E4AF CRC64; | | | |

Query Match 99.6%; Score 2832; DB 10; Length 548;
Best Local Similarity 99.6%; Pred. No. 1.9e-200;
Matches 546; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

| | | |
|----|-----|--|
| QY | 1 | MTTPTIANHPNFTHOSSSPLAFINTRTSIFPSSISKRNSVNCNGWTRCSVAKDYTP 60 |
| DB | 1 | MTTPTIANHPNFTHOSSSPLAFINTRTSIFPSSISKRNSVNCNGWTRCSVAKDYTP 60 |
| QY | 61 | SSAVYGGPAALDVCYVAGISGLCIAQVMSANPNIMVTEARRRAGNTTVRRDGIW 120 |
| DB | 61 | SSAVYGGPAALDVCYVAGISGLCIAQVMSANPNIMVTEARRRAGNTTVRRDGIW 120 |
| QY | 121 | EEGPNFQSDPDMITMAVDCGLKDDLVYGDPNAPRFVLMKGLRPVSKLTDLPFFDLMS 180 |
| DB | 121 | EEGPNFQSDPDMITMAVDCGLKDDLVYGDPNAPRFVLMKGLRPVSKLTDLPFFDLMS 180 |
| QY | 181 | IPGKIRAGFGPIGLRSPFGHESEYEQFVRNRLGGEVFERLIEPCSGYVYGDPSKLSMK 240 |

```
Db 181 IPKLRAGFGAIGLRPSPGHEESVEQFVRNRNLGGEVFERLIEPFCSGYVADDPKLSMK 240
Qy 241 AAFGKWKMLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKQGVGSFRKGLRMLPDA 300
Db 241 AAFGKWKMLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKQGVGSFRKGLRMLPDA 300
Qy 301 ISARLGSKLKLSWKLSITKSEKGYHLTYETPEGVVSLQSRISVMTVPVSVASNLIRPL 360
Db 301 ISARLGSKLKLSWKLSITKSEKGYHLTYETPEGVVSLQSRISVMTVPVSVASNLIRPL 360
Qy 361 SYAADAALSNFYPPVGVATISYPOEAIKRDRLVDGELKGQLHPRTOGVEITLGIYSS 420
Db 361 SYAADAALSNFYPPVGVATISYPOEAIKRDRLVDGELKGQLHPRTOGVEITLGIYSS 420
Qy 421 SLFPNRAKRGVLLNTYIGAKNPEILSKTESQLEVVDRDLRKMLIKPKADPLVVGVR 480
Db 421 SLFPNRAKRGVLLNTYIGAKNPEILSKTESQLEVVDRDLRKMLIKPKADPLVVGVR 480
Qy 481 VWPQAIPOFLVGHLDLTSTAKAAMNDNGLEGLFLGNYVSGVALGRCEGAYEVASEVTG 540
Db 481 VWPQAIPOFLVGHLDLTSTAKAAMNDNGLEGLFLGNYVSGVALGRCEGAYEVASEVTG 540
Qy 541 FLSTRYAK 548
Db 541 FLSTRYAK 548
```

```
RESULT 2
Q9SAX9 PRELIMINARY; PRT; 548 AA.
ID Q9SAX9;
AC 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PLASTIDAL PROTOPORPHYRINOGEN OXIDASE.
GN NTPPOX1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAISON NN;
RX PubMed=9608719;
RA Watanabe N., Che F., Iwano M., Takayama S., Nakano T., Yoshida S.,
RA Isogai A.;
RT "Molecular characterization of photomixotrophic cultured tobacco cells
RT resistant to protoporphyrinogen oxidase-inhibiting herbicides.";
RL Plant Physiol. 118:751-758(1998).
DR EMBL; AB020501; BAA34713.1; -.
DR InterPro; IPR000894; -.
DR Prodom; PD000290; -.
SQ SEQUENCE 548 AA; 59215 MW; AB51A513FD3CETBC CRC64;
```

```
Query Match 99.3%; Score 2824; DB 10; Length 548;
Best Local Similarity 99.3%; Pred. No. 7.6e-200;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 MTTTPPIAHNPITFHOSSSPLAFLNRTSPITFSSISKRSNVCNGMRTGCSVAKDYTPV 60
Db 1 MTTTPPIAHNPITFHOSSSPLAFLNRTSPITFSSISKRSNVCNGMRTGCSVAKDYTPV 60
Qy 61 SSAVDGPAELDCVIYAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
Db 61 SSAVDGPAELDCVIYAGISGLCIAQVMSANYPNLMVTEARDRAGNITTYVERDGYLW 120
Qy 121 EEGNRSFQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLTDLPEFDLMS 180
Db 121 EEGNRSFQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLTDLPEFDLMS 180
Qy 121 EEGNRSFQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLTDLPEFDLMS 180
Db 121 EEGNRSFQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLTDLPEFDLMS 180
```

```
Qy 181 IPKLRAGFGPIGLRSPSPGHEESVEQFVRNRNLGGEVFERLIEPFCSGYVADDPKLSMK 240
Db 181 IPKLRAGFGAIGLRPSPGHEESVEQFVRNRNLGGEVFERLIEPFCSGYVADDPKLSMK 240
Qy 241 AAFGKWKMLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKQGVGSFRKGLRMLPDA 300
Db 241 AAFGKWKMLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKQGVGSFRKGLRMLPDA 300
Qy 301 ISARLGSKLKLSWKLSITKSEKGYHLTYETPEGVVSLQSRISVMTVPVSVASNLIRPL 360
Db 301 ISARLGSKLKLSWKLSITKSEKGYHLTYETPEGVVSLQSRISVMTVPVSVASNLIRPL 360
Qy 361 SYAADAALSNFYPPVGVATISYPOEAIKRDRLVDGELKGQLHPRTOGVEITLGIYSS 420
Db 361 SYAADAALSNFYPPVGVATISYPOEAIKRDRLVDGELKGQLHPRTOGVEITLGIYSS 420
Qy 421 SLFPNRAKRGVLLNTYIGAKNPEILSKTESQLEVVDRDLRKMLIKPKADPLVVGVR 480
Db 421 SLFPNRAKRGVLLNTYIGAKNPEILSKTESQLEVVDRDLRKMLIKPKADPLVVGVR 480
Qy 481 VWPQAIPOFLVGHLDLTSTAKAAMNDNGLEGLFLGNYVSGVALGRCEGAYEVASEVTG 540
Db 481 VWPQAIPOFLVGHLDLTSTAKAAMNDNGLEGLFLGNYVSGVALGRCEGAYEVASEVTG 540
Qy 541 FLSTRYAK 548
Db 541 FLSTRYAK 548
```

```
RESULT 3
O64384 PRELIMINARY; PRT; 557 AA.
ID O64384;
AC 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OC NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BINTJE;
RA Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-
CC IX + H(2)O.
CC EMBL; AJ225107; CAA12400.1; -.
CC DR Mendel; 29307; Soltu;3076;29307.
CC OXidoreductase.
CC KW OXidoreductase.
SQ SEQUENCE 557 AA; 60463 MW; AE2B15B17E0B89F8 CRC64;
```

```
Query Match 90.2%; Score 2565.5; DB 10; Length 557;
Best Local Similarity 88.9%; Pred. No. 8.7e-181;
Matches 495; Conservative 21; Mismatches 32; Indels 9; Gaps 2;
```

```
Qy 1 MTTTPPIAHNPITFHQ-----SSSP-LAFLNRTSPITFSSISKRSNVCNGMRTGRC 51
Db 1 MTTTAAVNHPSITFTHSPSPSSSSSSPSFLFNNTNPIFYFSSKRSNVCNGMRTGRC 60
Qy 52 SVAKDYTPVSSAVDGPAAELDCVIYAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
Db 52 SVAKDYTPVSSAVDGPAAELDCVIYAGISGLCIAQVMSANYPNLMVTEARDRAGNIT 111
Qy 112 TYERDGYLWEEGNFSQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLT 171
Db 112 TYERDGYLWEEGNFSQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLT 171
Qy 121 TYERDGYLWEEGNFSQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLT 180
Db 121 TYERDGYLWEEGNFSQSPDPMLTMAVDGKLDLVLDGPNAPRVYLWKGKLRPVPKLT 180
Qy 172 DLPEFDLMSIPKLRAGFGPIGLRSPSPGHEESVEQFVRNRNLGGEVFERLIEPFCSGYV 231
Db 172 DLPEFDLMSIPKLRAGFGPIGLRSPSPGHEESVEQFVRNRNLGGEVFERLIEPFCSGYV 231
```

| Db | 161 | DLEPFDLMSITGKLRAGGALGLRPSPEGYESEVQVVRNRLCAEYERLIEEFCGSVYA | 240 |
|----------|---|---|--------------|
| QY | 232 | GDPSKLSMKAAPGAVKWMKLEETGSSIIIGTFEKAIKERSSTEPKAPRPLPKQQTGVSFR | 291 |
| Db | 241 | GDPSKLTLMKAAFGKWMKLEQGTGSSIIIGTFEKAIKERSSNKPPRDLPLPPKQQTGVSFR | 300 |
| QY | 292 | KGLRMLPAISARLGSKLKLSMKLSSTTKSEKGYHLTYETPEGVYSLQSRSLVMYPSY | 351 |
| Db | 301 | KGLRMLPAISERLGSKRSKLSMKLSSTTKSEKGYLLTYETPEGVYSLRFSRLVMYPSY | 360 |
| QY | 352 | VASNLIRPLSVAADALSNFYPPVGAVTISYPOEALIRDRERLVGDGELKGGOLHPRTQV | 411 |
| Db | 361 | VASNLIRPLSVAADALSNFYPPVGAVTISYPOEALIRDRERLVGDGELKGGOLHPRSQV | 420 |
| QY | 412 | ETLGTIVSSSLFPNRPAPRGVLLNTYIGATNTEIVSKTESQVLEAVDRDLRMLIKPKA | 471 |
| Db | 421 | ETLGTIVSSSLFPNRPAPRGVLLNTYIGATNTEIVSKTESQVLEAVDRDLRMLIKPKA | 480 |
| QY | 472 | QDPLVGVGRVWPQALPQFLVGHDLTSLTAKRAAMDNLBGLFLGAGNVSGVALGRCEGA | 531 |
| Db | 481 | QDPLVGVGRVWPQALPQFLVGHDLTSLTAKRALSDNGLDGLFLGAGNVSGVALGRCEGA | 540 |
| QY | 532 | YEAVSEVTGFLSRYAK 548 | |
| Db | 541 | YEAVSEVTGFLSRYAK 557 | |
| RESULT 4 | | | |
| Q9ZSIS | | | |
| ID | 09ZSIS | PRELIMINARY; | PRT; 545 AA. |
| AC | 09ZSIS | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Created) | | |
| DT | 01-MAY-1999 (TREMBLrel. 10, Last sequence update) | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last annotation update) | | |
| DE | TL5B16.1.13 PROTEIN. | | |
| GN | TL5B16.13. | | |
| OS | Arabidopsis thaliana (Mouse-ear cress). | | |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; | | |
| OC | Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; | | |
| OC | Brassicales; Brassicaceae; Arabidopsis. | | |
| OX | NCBI_TaxID=3702; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=CV. COLUMBIA; | | |
| RA | WASHU; | | |
| RT | "The A. thaliana Genome Sequencing Project."; | | |
| RL | Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases. | | |
| RN | [2] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=CV. COLUMBIA; | | |
| RA | Stoneking T., Smith R.; | | |
| RT | "The sequence of A. thaliana TL5B16."; | | |
| RL | Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases. | | |
| RN | [3] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=CV. COLUMBIA; | | |
| RA | Waterston R.; | | |
| RL | Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases. | | |
| DR | EMBL; AF1049419; AAC72870.1; - | | |
| DR | Mendel; 39261; Arab; 3076; 39261. | | |
| DR | InterPro; IPR000209; - | | |
| DR | InterPro; IPR001613; - | | |
| DR | PRINTS; PRO0757; AMINBOXDASEF. | | |
| DR | PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1. | | |
| SO | SEQUENCE 545 AA; 58598 MW; 6AB825691A01BE70 CRC64; | | |

| | | | | |
|---------------------------|--------|---------------------|------------|-------------|
| Query Match | 72.5% | Score 2061; | DB 10; | length 545; |
| Best Local Similarity | 74.3%; | Pred. No. 1,3e-143; | | |
| Matches 404; Conservative | 49; | Mismatches 73; | Indels 18; | Gaps 5 |

| | | | |
|----|-----|---|------|
| Db | 3 | LSLBRPTTQSLLPBPSKRN-LRLVNYKPLRRCVAGGPPYVSSKIEGGGTTITTDVCI | 61 |
| QY | 77 | VGAGISGLCIAQVMSANY-----PLNMYTEARDRAGNITTVERDGYLMEBGPNSFQSDP | 1322 |
| Db | 62 | YGGGISGLCIAQALATKHPDAAPNLIYTEANDRVGNITTEENGFLWMEBGPNSFQSDP | 1212 |
| QY | 133 | MILMAVDCGLKDDLVLDGPPNAPRFLYMKGKLRPVPSKLTLDLPEFDDLMSIPGKLRAGFGPI | 1922 |
| Db | 122 | MLTMVWDGSLKDDLVLDGPTAPRFLYLMGKLRPVPSKLTLDLPEFDDLMSIGGKIRAGFGAL | 1812 |
| QY | 193 | GLRBPSPBHEHSVEQFVARNINIGEVEFELILBPCSG-----VYVGDGDSKLMKAFAFG | 2444 |
| Db | 182 | GIRPBPBPRESBVEEFVARNINIGDEVEFELILBPCSGRIDKTSVYVGDGDSKLMKAFAFG | 2411 |
| QY | 245 | KWYKLEETGSGISIGTFFAKIRKERSSTPPAPDOPRLPKPKQGTGVSFAKGLRMLPDAISAR | 3042 |
| Db | 242 | KWYKLEQNGGSIIGTFFALIDERRKNAPAEADPRLPKRQGTGVSFFKGLRMLPEAISAR | 3012 |
| QY | 305 | LGSKLKLSMKTSSITKSEKGGHLLTETPEGVSLQSHSIVMTPVSYVASNILRLPSVAA | 3642 |
| Db | 302 | LGSKVKLSMKTSLGTLKLESBGVNLTYETPDDGLTASVQSKVMTVPBHVASGLRLPLESAA | 3612 |
| QY | 365 | ADALSNFPYPPYGVANTYITPOEALRDEFLVNGELAKRGQOLHPRQGVETLGTITYSSSLFP | 4222 |
| Db | 362 | ANALSKLTYPPVAAASISYPEKAEARTTECLIDGELKGFQOLHPRQGVETLGTITYSSSLFP | 4212 |
| QY | 425 | NNAPKRGVLLNTYIGGAANPELILSTESQOLEVYVDRDLRKLLIKPKKODPLVGVGRVWPQ | 4842 |
| Db | 422 | NNAPKRGVLLNTYIGGSINTGTILSKSEBELVEAVDRDLRKLLIKPNSTDPKLGLGRVWPQ | 4812 |
| QY | 485 | ALPQFLVHLDLTLSTAKAANDNGLBGLFGLGANNYSGVALRCRVEGAYEVAASEVTGFLSR | 5444 |
| Db | 482 | ALPQFLVGHFDLIDTAKSSLTSSGEGFLGANNYVAGVALRCRVEGAYETAIEVANNMSR | 5411 |
| QY | 545 | YAYK 548 | |
| Db | 542 | YAYK 545 | |

[illegible]

```
OY 77 VGAGISGLCIAQVMSANY-----PMLMTEARDRAGNITTVTERDGYLWEEGPNFSQSDP 132
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 80 VGAGISGLCIAQALATKTAASVPIVYTEARDRVGNISTYERDGYLWEEGPNFSQSDA 139
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 133 MLTMAVDGKLDLVLGDPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPGKLRAFGPI 192
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 140 MLTMVVDGKLDLVLGDPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPGKLRAFGAL 199
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 193 GLRPPPHESVGEQFVNRNLGGEYFELRLEPFGSGVYVGPSPSKLSMKAAGKWKLEET 252
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 200 GFRPSPPRESEVEEFVNRNLGDEVEFRLLEPFGSGVYVGPSPSKLSMKAAGKWKLEON 259
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 253 GGSIIIGTFFKAIRKERSPPKAPRDLPRKPGQVGSFRKGLRMLPDAISRLSKSLKS 312
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 260 GGSIVGAFKAIQDKKNOKPRPRDLPRKPGQVGSFRKGLRMLPDAISRLSKSLK 319
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 313 WKLSITSEKGYHLIYETEGVVSQSISIVMTVPSYVSNILRPLSVAADALSNEY 372
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 320 WKLSITSEKNGYHLIYETEGVVSQSISIVMTVPSYVSNILRPLSVAADALSKEY 379
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 373 YPVAVAITISYPOEAIREDRLVDGELKFGQLHPRTOGVEFLGTYSSSLFPNAPRGRV 432
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 380 YPVAVAISITPKDAIRADRLIDQLKFGQLHPRSQGVETLGTIYSSSLFPNAPRGRV 439
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 433 LLNLITGAKNPDLISTESQLEVEVDRDLRMLIKPAQDPLVGVYVWPAIQFLVG 492
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 440 LLNLITGATNPDLISTESQLEGEIVADVDRDLTMLIRDAEDPLTLGVYVWPAIQFLIG 499
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 493 HLDITLSTKKAAMNDNGLEGLFNGVYSGVALGRVCGAYVASEVTGELSRVAYK 548
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 500 HYDILDSAKAALSSGFGOMFLGNGVYSGVALGRVCGAYVASEVTGELSRVAYK 555
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
RESULT 6
O9M629 PRELIMINARY; PRT; 535 AA.
```

```
ID O9M629 AC O9M629;
AC O9M629;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE PROTOPHYRINOCEN IX OXIDASE.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. B73 INBRED;
RA Volrath S.L.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AF218052; AAF26417.1; -
SQ SEQUENCE 535 AA; 56835 MW; 9EEFB0E2973F919 CRC64;
```

Query Match 70.0%; Score 1989.5; DB 10; Length 535;
Best Local Similarity 72.3%; Pred. No. 2.3e-138;
Matches 394; Conservative 49; Mismatches 69; Indels 33; Gaps 6;

```
OY 17 SSSSPLAFNLRTSPFSSISKRNSVNCNGWTRCSYAKDYTVSSAVDGGPA----- 69
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 11 TAAAPL--LNGTR--IPALRLRGLSLVRC-----AAVAGGAABAPASTG 50
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 70 --AELDICYVAGISGLCIAQVMSANY--PMLMTEARDRAGNITTVTER--DGYLWEEG 123
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 51 ARLSADCVVGGISGLCTAQLATRGVGVIVTEARARGNITTVTERDEEYLMEEG 110
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 124 PNFQSPDPMITMAVDCGLDVLGDPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPG 183
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 111 PNFQSPDPMITMAVDCGLDVLGDPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPG 170
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 184 KLRAFGFPIGLRPPPHESVGEQFVNRNLGGEYFELRLEPFGSGVYVGPSPSKLSMKA 243
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

```

```
DB 171 KLRAAGLALGIRPPPHGRESVEEFVNRNLGAVEFRLLEPFGSGVYVADPSKLSMKA 230
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 244 GKWKLEETGGSITIGTFFKAIRKERSSTPKAPRDLPRKPGQVGSFRKGLRMLPDAIS 303
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 231 GKWVLEETGGSITIGTITIOERSKNPKPRDARLPKPGQVGSFRKGLRMLPDAIS 290
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 304 RLSSKTLSSITRSEKGYHLIYETEGVVSQSISIVMTVPSYVSNILRPLSVA 363
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 291 SLSSKTLSSITRSDKGYLLETPEGVVSQAQSWITLPSYVSNILRPLSSD 350
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 364 ADALSNFYPPYGAATISYPOEAIREDRLVDGELKFGQLHPRTOGVEFLGTYSSSLF 423
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 351 ADALSNFYPPYGAATISYPOEAIREDRLVDGELKFGQLHPRSQGVETLGTIYSSSLF 410
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 424 PNRAPGRVILLNWTGAKNPDLISTESQLEVEVDRDLRMLIKPAQDPLVGVYVWPA 483
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 411 PNRAPGRVILLNWTGAKNPDLISTESQLEVEVDRDLRMLIKPAQDPLVGVYVWPA 470
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 484 QALPFLVGHLDITLSTKKAAMNDNGLEGLFNGVYSGVALGRVCGAYVASEVTGEL 543
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 471 QALPFLVGHLDITLSTKKAAMNDNGLEGLFNGVYSGVALGRVCGAYVASEVTGEL 530
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 544 RYAYK 548
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 531 RYAYK 535
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

```

```
RESULT 7
O9LRI8 PRELIMINARY; PRT; 562 AA.
ID O9LRI8 AC O9LRI8;
AC O9LRI8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PROTOPHYRINOCEN OXIDASE (PROTOX-I).
GN SO-POXI.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. TONIC; TISSUE=LEAVES;
RA Che F.S., Watanabe N., Iwano M., Inokuchi H., Takayama S., Yoshida S.,
RA Isogai A.;
RT "Molecular characterization and subcellular localization of
RT protoporphyrinogen oxidase in spinach chloroplasts.";
RL Plant Physiol. 0:0-0(2000).
DR EMBL: AB029492; BAA96808.1; -
DR InterPro: IPR001613; -
DR InterPro: IPR001648; -
DR PRINTS: PR00757; AMINOXIDASEF.
DR ProDom: PD001272; -; 1.
SQ SEQUENCE 562 AA; 59929 MW; 638DCF6E63D259C4 CRC64;
```

Query Match 69.1%; Score 1966.5; DB 10; Length 562;
Best Local Similarity 71.1%; Pred. No. 1.2e-136;
Matches 381; Conservative 58; Mismatches 84; Indels 13; Gaps 5;

```
OY 26 NR-TSPFSSISKR--NSVNCNGWTRCSVA-----KDYTVSSAVD--GCPAAELDCVI 76
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 27 NRITLIPSSSLRRKRGSSITRSTISMSAAANYNQKNITNGVDGGGGVLDVCI 86
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 77 VGAGISGLCIAQVMSANYPNL---MYTEARDRAGNITTVTERDGYLWEEGPNFSQSDP 132
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 87 VGAGISGLCIAQALSTKYSNLSTNFIVTEAKDRVGNITTMADGGLWEEGPNFSQSDA 146
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
OY 133 MLTMAVDCGLDVLGDPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPGKLRAFGPI 192
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 147 VLTMAVDSGLKEELVLDGPNAPRFVLMGKLRPVPSKLTDLPEFDLMSIPGKLRAFGAL 206
||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

```

```

OY 193 GIRSPGHEESVQFVARNLGGVFERLLIEPCCSGVYVGDPSKLSMAAFGKWLKEET 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 207 GIRPSPHAEHSVEQFVRNRLGDEVERLIEPCCSGVYVGDPSKLSMAAFGKWLKEOK 266
OY 253 GGSIIIGFKAIKRRSSPPKAPRDLPKPKGVQVVSFRRGLRLPALISARLSKLS 312
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 267 GGSIIIGFKAIKRRSSPPKAPRDLPKPKGVQVVSFRRGLRLPALISARLSKLS 336
OY 313 WKLSITRSEKGYHLYEPEGVVSLQSRISVMTVPSYVANSILRPLSYAADAALSNFY 372
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 327 WLSGIKSNNGEYMLTYETPDGLVSRKSVMTVPSYVANSILRPLSDVAESLSKPH 386
OY 373 YRPVAVITVYPOAIRERLVDDELKGFGLHRTGVETLGTIYSSSLFPNRAKGRV 432
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 387 YRPVAVVSLSTPKKAIKRRSSPPKAPRDLPKPKGVQVVSFRRGLRLPALISARLSKLS 446
OY 433 LLANTIGGAKNPDLISKESOLVENVDRDLKMLIKKAPDPLVGVVRWPAIIPQFLVG 492
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 447 LLANTIGGAKNPDLISKESOLVENVDRDLKMLIKKAPDPLVGVVRWPAIIPQFLVG 506
OY 493 HLDPLSTAKAMNDNGLEGLFLGNYVSGVALGRCVGAATEVASEVTGELSRAYK 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 507 HFDLDAKKAALTDGGRKGLFLGNYVSGVALGRCVGAATEVASEVTGELSRAYK 562

```

```

RESULT 8
O92TA7 PRELIMINARY: PRT: 563 AA.
ID O92TA7
AC O92TA7
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
DE PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN PPX1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonads.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-621;
RX MEDLINE=99077317; PubMed=9862501;
RA Randalph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA Boynton J.E.;
RT "Isolation and characterization of a mutant protoporphyrinogen oxidase
RT gene from Chlamydomonas reinhardtii conferring resistance to porphyrin
RT herbicides."
RL Plant Mol. Biol. 38:839-858(1998).
DR EMBL: AF068635; AAC79685.1; -.
DR Mendel: 39099; Chlre:3076;39099.
SQ SEQUENCE 563 AA; 59802 MW; FC5E8FFC934CAF2 CRC64;

```

```

Query Match 53.8%; Score 1530.5; DB 10; Length 563;
Best Local Similarity 59.5%; Pred. No. 1.7e-104;
Matches 300; Conservative 73; Mismatches 114; Indels 17; Gaps 6;
OY 61 SSAVDGGPAE-----IDCVIVGAGISGLCTAQMVSANY--PMLMTEARDRAG 107
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 AAAATGAPTAGAGVAKTLDNVYDVIVYVGGGLSLVTGQALAAQHKQNLVTEARERVC 119
OY 108 GNITTVEDGYLMEGPRSPDPMLTMVDCGLKDDVLGDPNAPRFYLMGKLRPV 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 120 GNITSMGSDGYLMEGPRSPDPNMLQIADVSCGKDIYFGDPTARFYLMGKLRPV 179
OY 168 SKLIDLPEFDLMSIPGKIRAGFGIGL-RSPPGHEESVQFVARNLGGVFERLLIEPCC 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 SGL-DAFTFDLMSIPGKIRAGLGAIGLNGAMPFEESVQFIRNRLGDEVFFRLIEPCC 238
OY 227 SGVYVGDPSKLSMAAFGKWLKEETGSIIGTFKAIKRRSSPPKAPRDL-PKPKQ 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 239 SGVYVGDPSKLSMAAFGKWLKEETGSIIGTFKAIKRRSSPPKAPRDL-PKPKQ 298
OY 286 TVGSFRKGLRLPALISARLSKLSITRSEKGYHLYEPEGVVSLQSRISV 345
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 299 TVGSFRKGLRLPALISARLSKLSITRSEKGYHLYEPEGVVSLQSRISV 358
OY 346 MTVPSTVANSILRPLSYAADAALSNFYPPYGAVTISYPOEATIRDERLV-DGELKGFGL 404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 359 LTPSYVADVKKQAPPAEALGSPYPPVGAVTLSYPSAVREERKASDGSVPFGQL 418
OY 405 HPRDGVETLGTIYSSSLFPNRAKGRVLLNTYIGGAKNPDLISKESOLVENVDRDLK 464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 419 HPRDGVETLGTIYSSSLFPNRAKGRVLLNTYIGGAKNPDLISKESOLVENVDRDLK 478
OY 445 MLIKKAPDPLVGVVRWPAIIPQFLVGHLDPLSTAKAMNDNGLEGLFLGNYVSGVAL 524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 479 MLIKKAPDPLVGVVRWPAIIPQFLVGHLDPLSTAKAMNDNGLEGLFLGNYVSGVAL 538
OY 525 GRCVGAATEVASEVTGELSRAYK 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 539 GKVDEHGYESANNAKSVKAAVK 562

```

```

RESULT 9
O9TOP9 PRELIMINARY: PRT: 563 AA.
ID O9TOP9
AC O9TOP9
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
DE HERBICIDE-RESISTANT PROTOPORPHYRINOGEN OXIDASE PRECURSOR.
GN PPX1.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonads.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB-2674;
RX MEDLINE=99077317; PubMed=9862501;
RA Randalph-Anderson B.L., Sato R., Johnson A.M., Harris E.H.,
RA Hauser C.R., Oeda K., Ishige F., Nishio S., Gillham N.W.,
RA Boynton J.E.;
RT "Isolation and characterization of a mutant protoporphyrinogen oxidase
RT gene from Chlamydomonas reinhardtii conferring resistance to porphyrin
RT herbicides."
RL Plant Mol. Biol. 38:839-858(1998).
DR EMBL: AF030179; AAC79630.1; -.
FT TRANSIT PEPTIDE.
FT CHAIN 1 82 POTENTIAL.
FT 83 563 HERBICIDE-RESISTANT PROTOPORPHYRINOGEN
FT VARIANT 389 389 OXIDASE.
FT SEQUENCE 563 AA; 59834 MW; 074584EC935CA3AF CRC64;

```

```

Query Match 53.7%; Score 1527.5; DB 10; Length 563;
Best Local Similarity 59.3%; Pred. No. 2.8e-104;
Matches 299; Conservative 74; Mismatches 114; Indels 17; Gaps 6;
OY 61 SSAVDGGPAE-----IDCVIVGAGISGLCTAQMVSANY--PMLMTEARDRAG 107
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 AAAATGAPTAGAGVAKTLDNVYDVIVYVGGGLSLVTGQALAAQHKQNLVTEARERVC 119
OY 108 GNITTVEDGYLMEGPRSPDPMLTMVDCGLKDDVLGDPNAPRFYLMGKLRPV 167
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 120 GNITSMGSDGYLMEGPRSPDPNMLQIADVSCGKDIYFGDPTARFYLMGKLRPV 179
OY 168 SKLIDLPEFDLMSIPGKIRAGFGIGL-RSPPGHEESVQFVARNLGGVFERLLIEPCC 226
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 180 SGL-DAFTFDLMSIPGKIRAGLGAIGLNGAMPFEESVQFIRNRLGDEVFFRLIEPCC 238
OY 227 SGVYVGDPSKLSMAAFGKWLKEETGSIIGTFKAIKRRSSPPKAPRDL-PKPKQ 285
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

[illegible]

| RESULT | 10 |
|--------|---|
| Q9SLM5 | |
| ID | Q9SLM5 |
| AC | Q9SLM5 |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update) |
| DT | 01-MAR-2001 (TREMBLrel. 16, Last annotation update) |
| DE | PROTOPORPHYRINOGEN IX OXIDASE. |
| GN | HEMG. |
| OS | Glycine max (Soybean). |
| OC | Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; OC Fabales; Fabaceae; Papilionoideae; Glycyne. |
| OX | NCBI_TaxId=3847; |
| RP | [1] |
| RR | SEQUENCE FROM N.A. |
| RX | MEDLINE=20063348; PubMed=10594126; |
| RA | Kanjo N., Nishio S., Narita S., Oeda K., Inokuchi H.; |
| RT | Nucleotide sequence of a cDNA clone encoding protoporphyrinogen IX oxidase (Accession No. AB025102) from soybean (PG989-185).* |
| RL | Plant Physiol. 121:1383-1383(1999). |
| DR | EMBL: AB025102; BAA76348.1; -. |
| DR | InterPro: IPR00205; -. |
| DR | InterPro: IPR00759; -. |
| DR | InterPro: IPR002937; -. |
| DR | Pfam: PFO1593; Amino_oxidase; 1. |
| DR | PRINTS: PR00419; ADXRDPASE. |
| DR | SEQUENCE 502 AA: 55046 MW: 5908DEF6EADC2444 CRC64; |

| | | | | |
|---------------------------|--------|--------------------|------------|-------------|
| Query Match | 18.4% | Score 522.5; | DB 10; | Length 502; |
| Best Local Similarity | 29.1%; | Pred. No. 2.7e-30; | | |
| Matches 148; Conservative | 92; | Mismatches 225; | Indels 43; | Gaps 13; |

[illegible]

```

OY 292 KGLRMPDIAISARLG-SKIAKISAKLSITTSSEK-----GYHILYETPEGVASIQSMSIV 3454
Db 242 GGMQOTLDTLCKELKDKDLKNEKYLTLAAGHODSSSQMSWSTIASMSQSDVD--AVI 2999
OY 346 MPEVSIVASNIL-----RPLSAVAADALSNFPEYPCGAVTISYPOEALIRDLRVGEL 3589
Db 300 MTAPELVNVMADIKITKRGPFPLNF-----LPEVSIVPISWITTEGKEVVK-----RPL 3488
OY 399 KGRGQOLHP---RTQGYETLGIITLYSSSLFPPNRPKRGVLLNTIIGKAKNPBILSKTESOLV 4555
Db 349 EGGGVYVPEKQKNGIKTLGTLSMMFPRAASDLYLTFTTGCGQNEALQASTDEL 4081
OY 456 EYVDRLDKRLIKPKRAOPLVYGVAVMPDAILPQFLVGHLDLTSTAKAAMNDNGLEGFLG 5151
Db 409 KITSDRLKTL--GAEGEYTFVNFHFYMSKGPPLVGRNYGSVQIALDKIEKD--LPGFFFA 4666
OY 516 GNTVSGVALGRCVEGAETVYASVYVGFSL 543
Db 465 GNTKGLSVGKAIASGCAAADLVISYLN 492

```

| RESULT | 11 | | | |
|--------|--|--------------|------|---------|
| ID | 067814 | PRELIMINARY; | PRT; | 436 AA. |
| AC | 067814; | | | |
| DT | 01-AUG-1998 (TREMBLrel. 07, Created) | | | |
| DT | 01-AUG-1998 (TREMBLrel. 07, Last sequence update) | | | |
| DT | 01-JUN-2000 (TREMBLrel. 14, Last annotation update) | | | |
| DE | PHOTOPORPHYRINOGEN OXIDASE. | | | |
| GN | HEMG. | | | |
| OS | Aquifex aeolicus. | | | |
| OC | Bacteria; Aquificales; Aquificaceae; Aquifex. | | | |
| OX | NCBI_Taxid=63363; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RF | STRAIN=VF5; | | | |
| RX | MEDLINE=98196666; PubMed=937320; | | | |
| RA | Dackert G., Warren P.V., Gaasterland T., Young W.G., Ienox A.L., | | | |
| RA | Graham D.E., Overbeek R., Snead M.A., Keller M., Anjey M., Huber R., | | | |
| RA | Feldman R.A., Short J.M., Olson G.J., Swanson R.V.; | | | |
| RT | "The complete genome of the hyperthermophilic bacterium Aquifex | | | |
| RL | aeolicus."; | | | |
| RL | Nature 392:353-358(1998). | | | |
| DR | EMBL: AE000768; AAC07778.1; -. | | | |
| DR | InterPro: IPR000205; -. | | | |
| DR | InterPro: IPR002937; -. | | | |
| DR | Plam: PFO1593; Amino_oxidase; 1. | | | |
| QO | SEQUENCE 436 AA; 48987 MW; ED5F2B1BCDIDEEF7 CRC64; | | | |

| | | | | |
|-----------------------|-------------------|--------------------|------------|-------------|
| Query Match | 18.18; | Score 515; | DB 2; | Length 436; |
| Best Local Similarity | 29.98; | Pred. No. 7.7e-30; | | |
| Matches 144; | Conservative 103; | Mismatches 170; | Indels 64; | Gaps 16 |

```

QY 73 DCLVYVAGISGLCLIAQWMSANTYPMAMTLEARDRAGCNITTVBERDQYLMEECPNSQSPSDP 1322
Db 5 EYVYVAGISGLISSTAYRKLKRCBQVVDVYEDQDRIGGIHTHVKEKGLTEVEVGQITLADQE 64
QY 133 MLYMAVDCGLKJLDLVLAGDPNAP-REVLWAKGKLRVRVPSKLTLDLPFDLWMSITPEKLRAGFGP 1911
Db 65 VIFPLKEAGTIEP-VEASPSSKRYRIKRYKKRLLPLPMSVPEELKTLPLLLSTKTKLV-LTE 1212
QY 192 IGLRPSPCGHEE-VEGDFVFRNLNGEVEFERLLIEPSCGVYVDPGSKLSMKAAGKWKYL 249
Db 122 IFRK-----GVDEDISADPVEHGEELFWYVAPFISGVYAGDPEKLSLKHATPKLEA 1777
QY 250 EETFGSGLTGGTFRK-IKERSTPKRAPDRLPKPKGQYVGSFRKGLRMPLDAISRL--- 3051
Db 178 QKKYGGSLI-----KAIIKSKTA-----GPKKGLI-SFGESGIGELINAAQKLEVH 2211
QY 306 -GSKRLKSMKLSSTFKSEKGYHLYETFEPEGVAVSLQSSIMVTPSYVASNIIRPLSVA 3631

```

Db 222 TENVLRMR-KEEDFERLDVNG-----KKEVETKSVVAVSPATYSSTYLKEVSPS 269
QY 364 AADALSNFYPPVAVATISYPOEAIKIDERLVDGELKGFOLHPRTQGVETLTGTYSSSLF 423
Db 270 ASEEFDKIDYPPVVVNG-----VEGKFPKRVLLSSSGKRRILGAMFKSLF 318
QY 424 PNRAKGRVLLNTIGAKKNEILLSTESOLVEVVDRLKMLIKPRAODPLY--VGVRV 481
Db 319 PGRAPQOKELTTLVFGATREVIETSEETENIVERKEILT-----QIDCIDFMHVOK 373
QY 482 WPOAIPQIVGHLDTLSTAKAAMNDNGLEGLGNGVSVAGLCRCVEGAEVASEVTGF 541
Db 374 WKRAIPQITLGYDRFLMDAEMKD--YPGILFTGNMLYGVSTADICRASKVAQKVLTF 431
QY 542 L 542
Db 432 L 432

RESULT 12
Q9FYV8 PRELIMINARY; PRT; 544 AA.
AC Q9FYV8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DE 01-MAR-2001 (Tremblrel. 16, Last annotation update)
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
OC Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Volrath S.L.;
RL Submitted (MAY-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF273767; AAC00946.1; -
SQ SEQUENCE 544 AA; 59020 MW; 101CEFS6EB9A151B CRC64;

Query Match 16.1%; Score 459; DB 10; Length 544;
Best local Similarity 28.3%; Pred. No. 1.5e-25;
Matches 142; Conservative 88; Mismatches 226; Indels 46; Gaps 16;

QY 76 IVAGAGISGLCIAOVMSANYPNLMTVEARDRAGNITTVERRDGLMEGPNSPQSDPMLT 135
Db 50 VVAGAGVGLAAAYRLRQSGVNVTVFEADBRAGKIRTNSEGGFVMEGANTMEGEWAS 109
QY 136 MAV--DGLKXDDIVLGDPMNARFVLMKCKLRPVPSKLTDLFPFLDMSIPGKLRAFGPFI-- 192
Db 110 RLIDDLGLDQKQOYPNQHKRYIVKQDAPALIPSDPISLMKSSVLTSTKSLIALEFFPFLY 169
QY 193 ---GLRPPPGHE---ESVEQFVRNMGVEFERLIEPFCSGVYVGDPSKLSMKAAFGK 245
Db 170 KKANTRNSKAVSEHLSSESQSGFCEHNFGEVVDYFVDFVAGTSAGDPSLSIRAFRA 229
QY 246 VKMLEETGSGIIGCTFKAIKERSSTPKAPRDLPRPKQGVG--SEFKGLRMLPDAISAR 304
Db 230 LWNLEKRYGSAIVGAILSKIAKGDVKTTHDSSGRRNRNRFSPHGMOSLINALHNE 289
QY 305 LG-SKLKLSKLSITKS-----EKGQNHITVTPR--GVVSIQSR-----SIVTVASYYA 353
Db 290 VGDGNVAKLGEVLACTGPDGVALGRWSTISVDSKSGDKLASNOTFPAVITAP--L 346
QY 354 SNILRPL-----SVAADALSNFYPPVAVATISYPOEAIKIDERLVDGELKGFOLHP-- 406
Db 347 SNVRBMKFTTGAPVYLDLPFKMDYLPISLMTAFKKDDYK-----PLEGFGVLIIPK 400
QY 407 --RTQGVETLTGTYSSSLFPNRAKGRVLLNTIGAKKNEILLSTESOLVEVVDRLK 464
Db 401 EOKHGHKTIGTLTSSSMFPPDRAPDOYLYTTFVGGSHNRDLGAPTSILKQLVTSDLKK 460
QY 465 MLJKPKADPLVGVVRWVPOAIPQIFLVGHDLTSTAKAAMN--DNGLEGLFGLGNVYSGVA 523

Db 461 LL--GVGEQPTFYKHYWGNAP--LYGH-DVSSYLEALEKNEKNLPGFYVGNKNDGLA 515
QY 524 LGRVGAIEVASEVTGFLSRY 545
Db 516 VGSVIASGSKRADLAISYLESH 537

RESULT 13
Q9LYA7 PRELIMINARY; PRT; 501 AA.
AC Q9LYA7;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE PROTOPORPHRYNOGEN OXIDASE-LIKE PROTEIN.
GN F18022_10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL163817; CAB87761.1; -
DR InterPro: IPR000205; -
DR InterPro: IPR000759; -
DR PRINTS: PR00419; ADXKDTASE.
SQ SEQUENCE 501 AA; 54829 MW; 50CB8E1F3C0ED2F0 CRC64;

Query Match 16.0%; Score 456; DB 10; Length 501;
Best local Similarity 27.0%; Pred. No. 2.1e-25;
Matches 142; Conservative 95; Mismatches 194; Indels 94; Gaps 18;

QY 63 AVDGGPAELDCVIVAGISGLCIAOVMSANYPNLMTVEARDRAGNITTVERRDGLMEE 122
Db 13 AVSGKRVV-----VVGAGVGLAAAYKLLKSLNLTVEADRGVKGKLRVWNGILNDE 67
QY 123 GNSNSQPSDPMI--TMAYDCGLKDDLVLDGNPARRVLMKGLRPVPSKLTDLFPFLDMSI 181
Db 68 GANTMTAEPEVGSLLDDGLREKQOFPISQKRYIVRNGVPVMLPTNPIDLVYSSVLST 127
QY 182 PKLRAFGPFI-----GLRPPPGHESEVQFVRNMGVEFERLIEPFCSGVYVGDPSK 236
Db 128 QSKPQILLEPFLMKKSSVSDASAESEVSEFQNHQGVVDYLDIDFVGCTSAADPS 187
QY 237 LSMKAAFGKVMLEETGSGIIT-----GTFKAIKERSSTPKAPRDLPRPKQ 285
Db 188 LSMKHSFPLD--NSFSGIIVGAIKTFKAAGKGRSDTKSPGKKSGRS----- 236
QY 286 TVGSPRKGRLMDAI-----SARLSK--LKLS-----WKLSSITKSEKGVHL 328
Db 237 --FSFGKQMLIDPLTCLSKSLSHDETINLSKVLSTSYNSGROEMWSTLSCVSHNE----- 288
QY 329 TVETEGVVSLOSRSIVMTVPSYVANSI--TRPLSVAAA-----DALSNFYPPVAVT 380
Db 289 -----TQRN-----PHYDAAPLCNVKEMVVMGQPFQNLPELINTVPLSYLI 333
QY 381 ISYPOEAIKIDERLVDGELKGFOLHP--RTQGVETLTGTYSSSLFPNRAKGRVLLNTY 437
Db 334 TTFTEKRYK-----RPLEGFGVLIIPSKQKHGKFTLTGTLSSMMFPPDRSPSDVHLTTF 387
QY 438 IGGAKNPELLSTESOLVEVVDRLKMLIKPRAODPLYGVVRWVPOAIPQIFLVGHDLT 497
Db 388 IGGSRNOELAKASTIDELQVNTSDLRLL--GVGEQPVSNHYWVWRAKAPLYDSSY--DSV 444

QY 498 STAKAANDNGLEFLFGNVTYSGVALGRCEGAYEVAEVTGFL 542
DB 445 MEALDKM-ENLDLPGFFYAGNHRGGLSVGKSIASGCKADLVISYL 488

RESULT 14

Q9FMS9 PRELIMINARY; PRT; 547 AA.
AC Q9FMS9; 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
OS PROTOPORPHYRINOGEN IX OXIDASE.
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones."
RL DNA Res. 4:401-414(1997).
DR EMBL: AB007650; BAB08301.1; -
SQ SEQUENCE 547 AA; 60266 MW; 5474575A8B070A6E CRC64;

Query Match 16.0%; Score 456; DB 10; Length 547;
Best Local Similarity 27.0%; Pred. No. 2,3e-23;
Matches 142; Conservative 95; Mismatches 194; Indels 94; Gaps 18;

QY 63 AVDGPAAELDCVIVAGISGLCIAOVMSANYPNLMTYEDRAGCNITTYERDYLME 122
DB 59 AVSGKRAV-----VVGAGVSGLAAYKIKSLGNLNTYFADGRGKGLRSMQGLIMDE 113
QY 123 GPNFOPSPML-TMAVDCGLKDLVLGDPAAPRVLMKGLRPVSKLTDLPEDLMST 181
DB 114 GANMTAEAPVGLDLDLGRKQGFPIISQKRYIVANGVPVLPNPIELTVSSVLT 173
QY 182 PGKLRAGGPI-----GLRSPPGHESEVDFVNRNLGGEFPERLIRPFGSGVYVGPSPK 236
DB 174 QSKFOILLPEFLMKKSSKVSASAESEVSEFQRHFGQEVVDLIDPFVGGSADPDS 233
QY 237 LSMKAFGKVKMLEETGSI-----GGTFKAIKERSSTPKAPRDLPRKPGQ 285
DB 234 LSMKISFPDLM---NSFSIIVGAIRTFKAAGKGRPTKSPGTTKGSRS----- 282
QY 286 TVGSEFRKGLRMLPAI-----SARLGSK-LKLS-----WKLSTITKSEKGYHL 328
DB 283 --FSFKGGMQLPDLKSLSHDEINLDSKVLSTSYNGSRQEWMSLSCVSHNE----- 334
QY 329 TYEPREGVVSLSRSIVTVPSYASNI--LRPLSYAAA-----DALSNYYTPPGAVT 380
DB 335 -----TORON-----PHYDAAPLCNVKEMKVMKGGQFQDLPETINPLSLVLI 379
QY 381 ISYPOEAIROERLVDGLKFGOLHP---RTQGVETLGTITYSSSLFPNRAKGVLLNTY 437
DB 380 TTFYKEKK-----RPLEGFGVILPSKEQKHGKFTLGTLSMMFMFDRSRSDVHLVTTF 433
QY 438 IGGAKNPILKSTESQVLEVVVDRLKMLIRKPADPLVGVVRVPAIPQVLGHLDL 497
DB 434 IGGSRNQLAASDELKQVYTSDLQRL--GVEGEPSVNHYYRKAFFPLDSSY-DSV 490
QY 498 STAKAANDNGLEFLFGNVTYSGVALGRCEGAYEVAEVTGFL 542
DB 491 MEALDKM-ENLDLPGFFYAGNHRGGLSVGKSIASGCKADLVISYL 534

RESULT 15

AC 064385 PRELIMINARY; PRT; 404 AA.
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DE 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
OS PROTOPORPHYRINOGEN OXIDASE (EC 1.3.3.4).
OC Solanum tuberosum (potato).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. BINTJE;
RA Johnston D.J., Droz E., Rochaix J.D., Malnoe P.C.;
RL Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: PROTOPORPHYRINOGEN-IX + O(2) -> PROTOPORPHYRIN-IX + H(2)O.
DR EMBL: AJ225108; CAA12401.1; -
DR Mendel; 29309; Soltu;3077;29309.
DR InterPro; IPR000205; -
KW Oxidoreductase.
SQ SEQUENCE 404 AA; 44454 MW; A16C21D30AC46A34 CRC64;

Query Match 14.9%; Score 424; DB 10; Length 404;
Best Local Similarity 28.7%; Pred. No. 3,5e-23;
Matches 123; Conservative 83; Mismatches 152; Indels 70; Gaps 15;

QY 60 PSSAVDGPAAELDC-----VIVAGISGLCIAOVMSANYPNLMTYEDRAGCNITTYER 115
DB 3 PSAGED-----KONCKRAVIVAGVSGLAAYKIKHNLNTYFADGRGKGLRSLSQ 57
QY 116 DGLWEGNPSFOPSDPMLTMAVD-CGLKDLVLGDPAAPRVLMKGLRPVSKLTDLP 174
DB 58 DGLIMEGANMTSEBGVTFLLDSLGLREKQGFPLSQKRYIANGPTPLIPSPIDLI 117
QY 175 FPDLMSPGKLRAGFEPICLRSP-----PGHESEVDFVNRNLGGEFPERLIRPFGSGVY 230
DB 118 KSNFSLTSGSKLQMLEPRLMKNNKLLTKVSDHEVSGEFQRHFGKEVVDLIDPFVAGTC 177
QY 231 VGPSPKLSMKAAGKVKMLEETGSIIGSTFKA---IKERSSTPKAPRDLPRKPGQ 286
DB 178 GGDPPDLSMHLSPLEMLNLEKRGSVIYGAIRSKLSPIKEKQGP-----PKTSVKKRQ 232
QY 287 VGSFR--KGLRMLPAI-----SARLGSK-LKL-----SKLSTITKSEKGG 325
DB 233 RGSFSLGGMQTLTDAICDKLEDELRLNSRYLELSCSCGSDAIDSWISFASPHKROA 292
QY 326 YHLYTEPREGVVSLSRSIVTVPSYASNIILRPLSYAADAASNY-----YPPGAVTI 381
DB 293 EESSEFD-----AVIMTAAPLCVAKSM--KIAKKNPFLNFIPIVDVYVPLSVYIT 339
QY 382 SYPOEAIROERLVDGLKFGOLHP---RTQGVETLGTITYSSSLFPNRAKGVLLNTY 438
DB 340 TYKKEGVK-----PLEGFGVILPSQEQKHGKFTLGTLSMMFMFDRAPVNNVLYTTFV 393
QY 439 GGAKNPEI 446
DB 394 GGSNNREL 401

Search completed: July 3, 2001, 10:42:33
Job time: 571 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 3, 2001, 10:27:57 ; Search time 61.92 Seconds
(without alignments)
674.155 Million cell updates/sec

Title: US-09-508-418-2

Perfect score: 2844

Sequence: 1 MTTTPIANHPNIFTHQSSSS.....EGAYEVA5EVTGFLSRYAK 548

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: ptr1: *
2: ptr2: *
3: ptr3: *
4: ptr4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 2821 | 99.2 | 548 | 2 T04058 | protoporphyrinogen |
| 2 | 2565.5 | 90.2 | 557 | 2 T07116 | protoporphyrinogen |
| 3 | 2051 | 73.0 | 537 | 2 G85021 | protoporphyrinogen |
| 4 | 2065 | 72.5 | 545 | 2 T02005 | protoporphyrinogen |
| 5 | 513 | 18.1 | 436 | 2 B70473 | protoporphyrinogen |
| 6 | 505 | 17.8 | 504 | 2 T04076 | protoporphyrinogen |
| 7 | 485 | 17.1 | 470 | 2 D47045 | protoporphyrinogen |
| 8 | 456 | 16.0 | 501 | 2 T48595 | protoporphyrinogen |
| 9 | 424 | 14.9 | 404 | 2 T07121 | protoporphyrinogen |
| 10 | 407.5 | 14.3 | 467 | 2 D83800 | protoporphyrinogen |
| 11 | 368.5 | 13.0 | 477 | 2 S65684 | protoporphyrinogen |
| 12 | 363.5 | 12.8 | 462 | 2 A75435 | protoporphyrinogen |
| 13 | 363.5 | 12.8 | 477 | 2 A56449 | protoporphyrinogen |
| 14 | 362.5 | 12.7 | 477 | 2 S68367 | protoporphyrinogen |
| 15 | 317 | 11.1 | 468 | 2 T34679 | protoporphyrinogen |
| 16 | 302.5 | 10.6 | 490 | 2 T38088 | protoporphyrinogen |
| 17 | 254 | 8.9 | 435 | 2 F86601 | protoporphyrinogen |
| 18 | 254 | 8.9 | 435 | 2 A81518 | protoporphyrinogen |
| 19 | 249 | 8.8 | 435 | 2 D72023 | protoporphyrinogen |
| 20 | 231.5 | 8.1 | 424 | 2 D81739 | protoporphyrinogen |
| 21 | 221 | 7.8 | 426 | 2 F70869 | probable hemy - M |
| 22 | 211 | 7.4 | 424 | 2 A71476 | probable protoporp |
| 23 | 189 | 6.6 | 539 | 2 S50472 | hypothetical prote |
| 24 | 174.5 | 6.1 | 421 | 2 C75327 | oxidoreductase, FA |
| 25 | 169.5 | 6.0 | 441 | 2 A75361 | hypothetical prote |
| 26 | 153.5 | 5.4 | 416 | 2 T37023 | probable oxidoredu |
| 27 | 148 | 5.2 | 440 | 2 B71153 | hypothetical prote |
| 28 | 146 | 5.1 | 430 | 2 E84212 | hypothetical prote |
| 29 | 144 | 5.1 | 488 | 2 T47787 | hypothetical prote |

| | | | | | |
|----|-------|-----|-----|----------|--------------------|
| 30 | 143 | 5.0 | 490 | 2 A84861 | probable amine oxi |
| 31 | 141 | 5.0 | 448 | 2 H70947 | hypothetical prote |
| 32 | 139 | 4.9 | 570 | 2 S54134 | phytoene dehydroge |
| 33 | 136.5 | 4.8 | 467 | 2 T35307 | probable oxidoredu |
| 34 | 134 | 4.7 | 499 | 2 T51346 | monomine oxidase |
| 35 | 132.5 | 4.7 | 527 | 2 S03974 | amine oxidase (fla |
| 36 | 131 | 4.6 | 566 | 2 F71403 | hypothetical prote |
| 37 | 127 | 4.5 | 574 | 2 A96612 | hypothetical prote |
| 38 | 125.5 | 4.4 | 472 | 2 S74886 | phytoene dehydroge |
| 39 | 123.5 | 4.3 | 526 | 2 JT0528 | amine oxidase (fla |
| 40 | 123 | 4.3 | 582 | 2 S29314 | phytoene dehydroge |
| 41 | 122.5 | 4.3 | 489 | 2 A47259 | corticosteroid-bi |
| 42 | 122 | 4.3 | 471 | 2 S76290 | hypothetical prote |
| 43 | 121 | 4.3 | 404 | 2 A69467 | hypothetical prote |
| 44 | 119.5 | 4.2 | 472 | 2 T35936 | probable dehydroge |
| 45 | 119.5 | 4.2 | 478 | 2 A47693 | putrescine oxidase |

ALIGNMENTS

| | |
|---|--|
| RESULT 1 | |
| T04058 | protoporphyrinogen oxidase (EC 1.3.3.4) IX precursor, chloroplast - common tobacco |
| C:Species: Nicotiana tabacum (common tobacco) | |
| C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000 | |
| C:Accession: T04058 | |
| R:Rermonova, I.; Kruse, E.; Mock, H.P.; Grilmm, B. | |
| Proc. Natl. Acad. Sci. U.S.A. 94, 8895-8900, 1997 | |
| A>Title: Cloning and characterization of a plastidal and a mitochondrial isoform of t | |
| A:Reference number: 215186; MUID: 97385200 | |
| A:Accession: T04058 | |
| A>Status: preliminary; translated from GB/EMBL/DBJ | |
| A:Molecule type: mRNA | |
| A:Residues: 1-548 <LEB> | |
| A:Cross-references: EMBL:Y13465; NID:92370332; PIDN:CAA73865.1; PID:92370333 | |
| A:Experimental source: Strain SK1 | |
| C:Genetics: | |
| A:Gene: ppx1 | |
| C:Function: | |
| A:Pathway: tetrapyrrole synthesis | |
| C:Keywords: chloroplast; oxidoreductase | |
| F:1-50/Domain: transit peptide (chloroplast) #status predicted <TNP> | |
| F:51-548/Product: protoporphyrinogen oxidase IX #status predicted <MAT> | |
| Query Match | 99.2% Score 2821; DB 2; Length 548; |
| Best Local Similarity | 99.3% Pred. No. 2.1e-200; |
| Matches 544; Conservative 1; Mismatches 3; Indels 0; Gaps 0; | |
| QY 1 MTTTPIANHPNIFTHQSSSSPLAFINRTSFIPESSISKRSVNCNGWTRCSVAKDYTP 60 | |
| DB 1 MTTTPIANHPNIFTHQSSSSPLAFINRTSFIPESSISKRSVNCNGWTRCSVAKDYTP 60 | |
| QY 61 SSAVGGPRAELDCYIVAGISGLCIAQVMSANYNLAVTARBRAGNITTVREDGLW 120 | |
| DB 61 SSAVGGPRAELDCYIVAGISGLCIAQVMSANYNLAVTARBRAGNITTVREDGLW 120 | |
| QY 121 EEGPNSFOPSDPMLTMAVDCGLKDDLVGDPNAPFVLMKSLRVPKLTDLPEFDLMS 180 | |
| DB 121 EEGPNSFOPSDPMLTMAVDCGLKDDLVGDPNAPFVLMKSLRVPKLTDLPEFDLMS 180 | |
| QY 181 IEGKLRAFGPIGLRPPSPGHEEVEQFVRNRLGGEVERLIEPFCGIVYGDPSKLSMK 240 | |
| DB 181 IEGKLRAFGPIGLRPPSPGHEEVEQFVRNRLGGEVERLIEPFCGIVYGDPSKLSMK 240 | |
| QY 241 AAFGVMKLEETGGSIIIGTFPAIKERSSTPAKAPDPLPKPGQTVGSFRKGLBMLPDA 300 | |
| DB 241 AAFGVMKLEETGGSIIIGTFPAIKERSSTPAKAPDPLPKPGQTVGSFRKGLBMLPDA 300 | |
| QY 301 ISARLGSKLKLSMKLSITKSEKGYHLTYETPEGVSLQSKSIYMPVPSVANSILRPL 360 | |
| DB 301 ISARLGSKLKLSMKLSITKSEKGYHLTYETPEGVSLQSKSIYMPVPSVANSILRPL 360 | |

| | | | | | | | | |
|-----------------------|--|---|---------------------|--------|-------------|-------------------------|-----------|-----|
| QY | 361 | SVAAADALSNFYYPVGA | AVITISYPOE | AI | RDE | RLVDGELKFGQLHPTGCVETLGT | YSS | 420 |
| Db | 361 | SVAAADALSNFYTPVGA | VITITYPQEA | I | RDE | RLVDGELKFGQLHPTGCVETLGT | YSS | 420 |
| QY | 421 | SLPPNRAKRGVLLN | YTGAKNPEILSKTESQ | LVEV | VD | DLKRMILTKPKAODPLVYGV | R | 480 |
| Db | 421 | SLPPNRAKRGVLLN | YTGAKNPEILSKTESQ | LVEV | VD | DLKRMILTKPKAODPLVYGV | R | 480 |
| QY | 481 | VWQQAIPQELVGH | LDLSTKKAAMNDNG | EGLE | ELG | CNYVSGVALGR | CVESAEYAS | 540 |
| Db | 481 | VWQQAIPQELVGH | LDLSTKKAAMNDNG | EGLE | ELG | CNYVSGVALGR | CVESAEYAS | 540 |
| QY | 541 | FLSRVAYK | 548 | | | | | |
| Db | 541 | FLSRVAYK | 548 | | | | | |
| RESULT | 2 | | | | | | | |
| | 707116 | | | | | | | |
| | protoperchlorophytinogen oxidase (EC 1.3.3.4) IX, chloroplast - potato | | | | | | | |
| | C:Species: Solanum tuberosum (potato) | | | | | | | |
| | C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999 | | | | | | | |
| | C:Accession: T07116 | | | | | | | |
| | R:Johnston, D.J. | | | | | | | |
| | submitted to the EMBL Data Library, April 1998 | | | | | | | |
| | A:Reference number: Z15932 | | | | | | | |
| | A:Accession: T07116 | | | | | | | |
| | A:Status: translated from GB/EMBL/DBJ | | | | | | | |
| | A:Molecule type: mRNA | | | | | | | |
| | A:Residues: 1-557 <CDS> | | | | | | | |
| | A:Cross-references: EMBL:AJ225107; NID:e1289157; PIDN:CAA12400.1; PID:e1289158 | | | | | | | |
| | A:Experimental source: cv. Bintje | | | | | | | |
| | C:Genetics: | | | | | | | |
| | A:Genome: nuclear | | | | | | | |
| | C:function: | | | | | | | |
| | A:pathway: tetrapyrrole biosynthesis | | | | | | | |
| | C:keywords: chloroplast; oxidoreductase | | | | | | | |
| Query Match | | 90.2%; | Score 2565.5; | DB 2; | Length 557; | | | |
| Best Local Similarity | | 88.9%; | Pred. No. 1.5e-181; | | | | | |
| Matches 495; | Conservative 21; | Mismatches 32; | Indels 9; | Gaps 2 | | | | |
| QY | 1 | MTTPIAHNPNIPTHQ-----SSSSP-LAFLNRTSFIPSSISKRSNVNCGMRTRC | 51 | | | | | |
| Db | 1 | MTTAVAHNPISITHTSPLRSPSSSSSSSFLEPLNTNTEIPFSTIKRSNVNCGMRTRC | 60 | | | | | |
| QY | 52 | SVAKDYTVSSAVDGGPAAELDCVITGAGISGICIAQVMSANPNLMVTEARDACGNT | 111 | | | | | |
| Db | 61 | SVAKDYTVPPSEVDGNQPELDCVVVGAGISGICIAKVISANPNLMVTEARDACGNT | 120 | | | | | |
| QY | 112 | TYVERDGYLMEEGNSRQSPDDPMITAAVDCGLKDDLVLDGPNAPREYLMWGLKLRPVSKLT | 171 | | | | | |
| Db | 121 | TYVERDGYLMEEGNSRQSPDDPMITAAVDCGLKDDLVLDGPNAPREYLMWGLKLRPVSKLT | 180 | | | | | |
| QY | 172 | DLPEFLMISIPGKLIRAGFGFPIGLRPPGHEESVEQGFVRNNGEVEERLIEPFCGGVY | 231 | | | | | |
| Db | 181 | DLPEFLMISIPGKLIRAGFGFPIGLRPPGHEESVEQGFVRNNGEVEERLIEPFCGGVY | 240 | | | | | |
| QY | 232 | GDPSKLSMKAAGCKYWKLEETGSGIIGTFPKAIKERSSTPKAPRDRPLPKPGQTVGSFR | 291 | | | | | |
| Db | 241 | GDPSKLSMKAAGCKYWKLEETGSGIIGTFPKAIKERSSNPKPRDRPLTPKPGQTVGSFR | 300 | | | | | |
| QY | 292 | KGLRMPLDAISARLGSKLMSKSSITTSSEGGYHLTETPEGVYSLORSISIVMTVPSY | 351 | | | | | |
| Db | 301 | KGLRMPLDAICERLSKSVLMSKSSITTSSEGGYHLTETPEGVYSLORSISIVMTVPSY | 360 | | | | | |
| QY | 352 | VASNLIRPLSVAAADALSNFYYPVGA | VITISYPOE | AI | RDE | RLVDGELKFGQLHPTGCV | 411 | |
| Db | 361 | VASNLIRPLSVAAADALSNFYYPVGA | VITITYPQEA | I | RDE | RLVDGELKFGQLHPTGCV | 420 | |
| QY | 412 | ETLGTITYSSTLFPNRAKRGVLLN | YTGAKNPEILSKTESQ | LVEV | VD | DLKRMILTKPKA | 471 | |

```

Db      421  ETLGTTTSSSLFPNNAPNCRVLLNLTIGATNTTLYSKTESQLYEAVDRDLRKMLIKFKA 480
QY      472  QDPLVGVVRVWQAIPLQGLVGHLDFTLSTAKAAMNDNGEGLFLGGNYVSGVALGRCEGA 531
      ||| :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      481  QDPFVTGVVRWQAIPLQGLVGHLDFTLGTAKTALSDNGLDGLFLGGNYVSGVALGRCEGA 540
QY      532  YEVASEVTGFLSRVAYK 548
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      541  YEIASEVTGFLSQVAYK 557

RESULT      3
G85021
protoporphyrinogen oxidase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: G85021
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory, 1999
A:title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: AB5001; MUID:20083488
A:Accession: G85021
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <STD>
A:Cross-references: GB:NC_001268; NID:g7268212; PIDN:CAB77739.1; GSPDB:GN00140
C:Genetics:
A:gene: AT4g01690
A:map position: 4

```

| Query Match | 73.0% | Score 2075 | DB 2 | Length 537 |
|-----------------------|-----------------|--|-----------|------------|
| Best Local Similarity | 75.4% | Pred. No. 2.4e-15 | | |
| Matches 404 | Conservative 49 | Mismatches 73 | Indels 10 | Gaps |
| QY | 22 | LAFINRTSFIPEFSISKRNNSVNCNGW---RTQCASAKDYTPVSSAVDGGCAAEU---DCVY | 76 | |
| Db | 3 | LSLIRPTTGGSLPEFSKPN-LRNIYKPRLRKCSVAGGPTVGSKIREGGCTITTTDCVY | 61 | |
| QY | 77 | VGAGISGLCIAQVMSANY-----PNIMVTEARDRAKGNITTYERGQYLMEEGPNSPQSP | 132 | |
| Db | 62 | VGGISGLCIAQALAKRHPDAPNLIYEAQRKQVGNITTEKNGFLMEEGPNSPQSP | 121 | |
| QY | 133 | MLTMVNDGGLKLDLVIGDPNAPRFVLMGKGLRPVPSKTLDPFPFDLMSIGKLRAGEGP | 192 | |
| Db | 122 | MLTMVNDGGLKLDLVIGDPTAPRFVLMGKLRPVPSKTLDPFPFDLMSIGKLRAGEGAL | 181 | |
| QY | 193 | GLRSPPGHEESVEQFVRNRLGGEFERLIPEFCGQVYVGPSPKLSMKAAPGVYMLLEET | 252 | |
| Db | 182 | GIRSPPGRESVEEYVVRNRLDGEYFERLIEFPCSGYAGDPSKLSMKAAPGVYMLLEON | 241 | |
| QY | 253 | GGSLIGTFKAIKERSSTPRADPRLPKPGQVTSFRKGLRMLPDAISARLSKULS | 312 | |
| Db | 242 | GGSLIGTFKAIERKNAPKARDPRLPKPGQVTSFRKGLRMLPEAISARLSKVLKS | 301 | |
| QY | 313 | WKLSSITKSKGQYHLTYETPEGCVSLOSRSIVMTVPSYVANSILRPLSLAADAALSNTY | 372 | |
| Db | 302 | WKLSSITKLSGCGYHLTYETPDGLVSYSKSVMTVPSHVASGLRPLSLASANAALSKLY | 361 | |
| QY | 373 | YPPYGAVTISPEALIRDELVDGELKGFQGHLPQTQGVETLGTIVSSSLFPPRAKGVY | 432 | |
| Db | 362 | YPPYAAVSISYKFAIRTECLIDGELKGFQGHLPQTQGVETLGTIVSSSLFPPRAKPGNI | 421 | |
| QY | 433 | LLNLYIGKANKPELISKTESQSLVEVANDLKKMLIKRAQDPLVYGVRWVQAIPQFLVG | 492 | |
| Db | 422 | LLNLYIGGSYINGILSKSGELVEAVDDLKMLIKRPNSTDPKLKGVRWVQAIPQFLVG | 481 | |
| QY | 493 | HHDTLSTAKAAMNDGLBGLTGSNYTSGVALGRCEBGAETVASEVTGLSTRAYK | 548 | |
| Db | 482 | HHDTLSTAKASSSTSSGEGTIGSYVGVAGVALGCEVGAETVIEVNNFESRAYK | 537 | |

```

RESULT      4
T02005
protoporphyrinogen oxidase (EC 1.3.3.4) - Arabidopsis thaliana
N:Alternate names: protein T15B16.13
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 26-Feb-1999 #sequence_revision 26-Feb-1999 #text_change 21-Jul-2000
C:Accession: T02005; J05488
R:Stonking, T.; Smith, R.
submitted to the EMBL Data Library, November 1998
A:Description: The sequence of A. thaliana T15B16.
A:Reference number: Z14488
A:Accession: T02005
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-545 <STO>
A:Cross-references: EMBL:AF104919; NID:g3859590; PID:g3859604
A:Experimental source: cultivar Columbia
R:Narita, S.; Tanaka, R.; Ito, T.; Okada, K.; Taketani, S.; Inokuchi, H.
Gene 182, 169-175, 1996
A:Title: Molecular cloning and characterization of a cDNA that encodes protoporphyrinogen
A:Reference number: J05488; MUID:97136707
A:Accession: J05488
A:Molecule type: mRNA
A:Residues: 1-217, 226-545 <NAR>
A:Cross-references: DDBJ:D83139; NID:g118306; PIDD:BA11820.1; PID:g1877018
C:Genetics:
A:Map position: 4
A:Introns: 126/3; 188/3; 225/1; 274/2; 301/2; 358/3; 412/1; 446/3
A:Note: T15B16.13
C:Function:
A:Description: catalyzes the removal of six electrons from protoporphyrinogen to generate
A:Pathway: chlorophyll biosynthetis; heme biosynthetis
C:Keywords: oxidoreductase

Query Match      72.5%; Score 2061; DB 2; Length 545;
Best Local Similarity 74.3%; Pred. No. 2.6e-144;
Matches 404; Conservative 49; Mismatches 73; Indels 18; Gaps 5;

0Y 22 TAFRTSFIPSSISKRNVNCGM---RTRCSVAKYTPSSAVDGPAAEL--DCVI 76
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 3 LSLRPTQSLPSRSKR-ELRNVKPLRLCSVAGGPYSSKIREGGGTTITDVI 61

0Y 77 VGAGISGICIAQVMSANY---PNLWTEARDRAGNITTVRDGYLWEEGPNSEFQSDP 132
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 62 VGGISGICIAQALATKHPDAPNLIVTEAKDRVGNITIRENGFLWEEGPNSEFQSDP 121

0Y 133 MTTMAVDCGLKDDVLVGPNAFVLMKGLRPVSKLTDLFFFLMSIPGKLKRGFGPI 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 122 MTTMVDVSLKDDVLVGDPTAFVLMNGKLRPVSKLTDLFFFLMSIGKIRAGFGAL 181

0Y 193 GLRPPSPGHEEVEQFVRNLGGEVFERLIEPFCG-----YYVGDPSTLSKKAAG 244
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 182 GLRPPSPGHEEVEERVRNLGDEVEFERLIEPFCGSRVIDKSCYAGDPSTLSKKAAG 241

0Y 245 KWKLEETGGSIIGGTFKAIKERSSTPKAPRDPRLPKPGQTVGSFRKGLMPLDAISAR 304
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 242 KWKLEQNGNSIIGGTFKAIQERKNAPKAERDPRLPKPGQTVGSPRGKLRMLPEAISAR 301

0Y 305 LGSKLKLSKLSITKSEGGYHLTYETPEGVVISLQNSIVTVPSTYASNLRLPLSAA 364
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 302 LGSKYLKLSKLSIGITKLESGYNLTETPDGLVSYOSKSVVTVPSHVASGLRLPLESA 361

0Y 365 ALALSFYPPYGAWTISYPOEAIQERIVDGLKGFQGLHRTGQVEFLGTVSSSLFP 424
   | : | : | | | | | | | | | | | | | | | | | | | | | | | |
DB 362 ALALSKLYPPAAVATISYPKFAIRTECLIDELKGFQGLHRTGQVEFLGTVSSSLFP 421

0Y 425 NRAPGRVLLNYIGAKAPPELSTESQLEVEVVDRLKMLIKPAQDPLVVGVRWFO 484
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 422 NRAPGRILLNYIGSTNTGLTSLKSEGLVEAVVRDLKMLIKRNSDPLKLGVRWFO 481

0Y 485 AIPQPLVGHLDPLTSPAKAAMDNGLEGLFVGNNYSGVALRCVCGAVVASEVTGFLSR 544
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

| | | | |
|---|--|--|-----|
| Db | 482 | AIRPQVLGHFPIIDLRKASSTLSSGEGLEFLGGNGVYACALRCVEAGVENALEVNMMR | 541 |
| OY | 545 | YAYK 548 | |
| Db | 542 | YAYK 545 | |
| RESULT | 5 | | |
| B70473 | | protoporphyrogen oxidase - Aquifex aeolicus | |
| C:Species: | | Aquifex aeolicus | |
| C:Date: | | 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999 | |
| C:Accession: | | B70473 | |
| R:Deckerl, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; | | | |
| V. | | Nature 392, 353-358, 1998 | |
| A:Title: | | The complete genome of the hyperthermophilic bacterium Aquifex aeolicus. | |
| A:Reference number: | | A70300; MUID:98196666 | |
| A:Accession: | | B70473 | |
| A>Status: | | Preliminary; nucleic acid sequence not shown; translation not shown | |
| A:Molecule type: | | DNA | |
| A:Residues: | | 1-436 <ACF> | |
| A:Cross-references: | | GB:A0000768; NID:g2984249; PIDN:AAC07778.1; PID:g2984251; GB:A000 | |
| C:Experimental source: | | strain VF5 | |
| C:Genetics: | | | |
| A:Gene: | | hemG | |
| Query Match | 18.1%; Score 515; DB 2; Length 436; | | |
| Best Local Similarity | 29.9%; Pred. No. 2.5e-30; | | |
| Matches | 144; Conservative 103; Mismatches 170; Indels 64; Gaps 16; | | |
| OY | 73 | DCVIYGAGISGLCIAQVMSANYPNLMATTEARDRAGGNITTYERDGYLWEEGPSNFOPSDP | 132 |
| Db | 5 | EVVVGAGISGLSTAYRLKKEGVDVYVEKDDRIIGTLHTFKKEKGYLEFGAQTLLADQE | 64 |
| OY | 133 | MLTMAVDCGLKDVLGVDPNAP-REVLMKGRIPPSKLTDLPFDLDLSIFGKLRAAGRP | 191 |
| Db | 65 | VIDFLKEAIEP--VEASPSSKYRYIKRKRLIPLMPSPVEFLKTPTLSLTKLV-LTE | 121 |
| OY | 192 | IGLRSPSGHEH--SVEDPVRRNLGGEVFERLIIEPGCGVYVGDPSKLSMKAAGKWKYL | 249 |
| Db | 122 | IHKR-----GVDEDISTADFVRHFGEELPLNYAVPFISGYAGADEPKLSLNHAIPTLEYA | 177 |
| OY | 250 | EETGSIIIGTFTKA-IKERKSTPKARPDRPLPKGTQVTSFRKGLRMAPPAINARL--- | 305 |
| Db | 178 | OKKYSGLI---KAFIKETA-----GPRGKLI-SFGESLGILNALAQKLEVH | 221 |
| OY | 306 | --GSKIKLSWKLSSITTKSEKGYHLITYETPEGSVLSQRSISLYMYPPSVASNILRPLSYA | 363 |
| Db | 222 | TENVVLRM-RKEDEFRLDVRG-----KKVERKSYYVASPATSYSLKEVSFS | 269 |
| OY | 364 | ABDALSNFYPPVGAVTISYPEAIRDERLVNDGELKRGGOJHPRTQGETVETLTGISSLF | 423 |
| Db | 270 | ASEEPDKIDYPVVVVVNVG-----VEGKTPPKYRLSSSEGKRRIICAMPMSKLF | 318 |
| OY | 424 | PNRAEGRVLLNTYIGAANKPELLKTESQILEVVDRLRKMLIKPKAADPLY-VGVRV | 481 |
| Db | 319 | PGRAPQGRKLLLVFLFGAGNDREVIELSEELINUYRELKELL-----QIDCIDPMHYQX | 373 |
| OY | 482 | WPQAIPQFLVGLDPLSLTAAKAMNDGLEGLFLGSNNYSVAALGCVCAGATEVASEVTGF | 541 |
| Db | 374 | MKRAIPQYTLTGDFRLNLAOEWEKD--YFGLEFLTGMNLGYGSTADCIIRASKRVAAQKVLSF | 431 |
| OY | 542 | L 542 | |
| Db | 432 | L 432 | |
| RESULT | 6 | | |
| T04076 | | protoporphyrogen oxidase (EC 1.3.3.4) IX, mitochondrial [validated] - common tobacco | |
| C:Species: | | Nicotiana tabacum (common tobacco) | |

148595
 protoporphyrinogen oxidase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F18022.10
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48595
 R:Bayan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 424493
 A:Accession: T48595
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-501 <BEV>
 A:Cross-references: EMBL:DB
 A:Experimental source: cultivar Columbia; BAC clone F18022
 C:Genetics:
 A:Map position: 5
 A:Features: 10/1; 27/2; 72/3; 94/3; 111/3; 130/3; 155/2; 167/3; 189/3; 199/3; 244/3; 298/
 A:Note: F18022.10

Query Match 16.0%; Score 456; DB 2; Length 501;
 Best Local Similarity 27.0%; Pred No 7e-26; Mismatches 194; Indels 94; Gaps 18;
 Matches 142; Conservative 95; Mismatches 194; Indels 94; Gaps 18;
 QY 63 AVDCGPAALDCVYAGISGLCTIAQVMSANTPNLMTVEARDRAGNTTVERDGLTME 122
 DB 13 AVSGKRVA-----VVGAGVSGLAAYKLRKGLNVTVEADRGKLRKNGMLTME 67
 QY 123 GPNSPQSDPML-TMAVDCGILKDLVLDGPNAPRFVLMKGLRPVSKLTDLPFDLMSI 181
 DB 68 GANTMTAEPEVSGILDLGLRERKQFPISOKRKYIVNGVPVLMPTNPELTVTSSVLT 127
 QY 182 PGKLAGRGP-----GLRSPGHESEYQFRRNLGCEFERLIPPCSGVYGDPSK 236
 DB 128 OSKFOILLPEFLMKKSSVDASAESEVSEFFQRHFOEVDYLDLPFGVGSADPSK 187
 QY 237 LSKKAAPFGKVKLEETGSGII-----GGTFKAIKERSGTPKAPDPPLPKPKQ 285
 DB 188 LSKKHSFPLM-----NSFGSIIVGAIRTKFAAGKSGKSTKSSPTKKGSGS----- 236
 QY 286 TVGSPFKGLRMLPDAI-----SARLGSK-LKLS-----WKLSTIKSEKGYHL 328
 DB 237 --FSFRGGQILPTLCKLSHDEINLDSKVLSTYNSGROENWSLSCVSHNE----- 288
 QY 329 TVETEGVVSIGRSIYMTVPSTYASNI--LRPLSYAA-----DALSNFYTPVGAVT 380
 DB 289 -----TORQN-----PHYDAAPLCNVKEMKVGCGQFQINFLPEINMPLSVLI 333
 QY 381 ISYPOAIRDERLVDGLKGFQLHP---RTQGVETLGTIYSSSLFPNRAPKRVLILNY 437
 DB 334 TTFTRKRYK-----RPLEGFGVILIRSKQKGFRTLSTSSMMPDRSPDVHLXTTF 387
 QY 438 IGGANPELSTESQVLEVYDRLKMLIKRQADPLVGVVWPAIQFLVGLIDL 497
 DB 388 IGSRRQELAKASDELKQVYSDLRLL--GVESEPVSVHYYWRKAPLYDSY--DSV 444
 QY 498 STAKAMNDNGLEGLFLGNGVSGVALGRCVGAATEVASEVTGFL 542
 DB 445 MEALDKM-ENDLPGFFYAGNHRLGLSVGKSIAGCKADLVISYL 488

RESULT 9
 T07121
 protoporphyrinogen oxidase (EC 1.3.3.4) IX, mitochondrial - potato (fragment)
 C:Species: Solanum tuberosum (potato)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
 C:Accession: T07121
 R:Johnston, D.J.
 submitted to the EMBL Data Library, April 1999
 A:Reference number: Z15392
 A:Accession: T07121
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
 A:Residues: 1-404 <JOH>
 A:Cross-references: EMBL:AJ225108; NID:e1289159; PIDN:CAA12401.1; PID:e1289160
 A:Experimental source: cv. Bintje
 C:Genetics:
 A:Genome: nuclear
 C:Function:
 A:Pathway: tetrapyrrole synthesis
 C:Keywords: mitochondrion; oxidoreductase

Query Match 14.9%; Score 424; DB 2; Length 404;
 Best Local Similarity 28.7%; Pred No 1.2e-23;
 Matches 123; Conservative 83; Mismatches 152; Indels 70; Gaps 15;
 QY 60 PSSAVDGPAAEIDC-----VIYAGISGLCTIAQVMSANTPNLMTVEARDRAGNTTVER 115
 DB 3 PSAGED-----KONCPKRVAVIAGVSGIAAAYKLIHGLNVTVEADRGKLRKLSIQ 57
 QY 116 DGYLMERGPNSPQSDPMLTMAVD--CGIKDOLVLDGPNAPRFVLMKGLRPVSKLTDLP 174
 DB 58 DGLIMDEGANTMTESGQVTLFLDSIGRERKQFPISOKRKYIARNGTPPLIPSPIDL 117
 QY 175 FFDLMSIPGKLRAGRPICLRSP---PGHEEVLDQFYRRNLGCEFERLIPPCSGY 230
 DB 118 KSNFLSTGSKLQMLFPLLMKNNKLTQYSDHESEVSGFFQRHGRKQVYLDLPVAVGTC 177
 QY 231 VGDPKLSMKKAPFGKVKLEETGSGIIGTFKA-----IKERSTPKAPDPRLPKPGQT 286
 DB 178 GGDSDSLSMHLSPELMLNLEKRGSVIVGAIRSKLSPEKKQGP---PRTSVNKRQ 232
 QY 287 VGSFR--KGLRMLPDAI-----SARLGSK-LKLS-----WKLSTIKSEKGG 325
 DB 233 RGSFSGFGQQLTDAICDKDELRLNLSVLELSCSGSADISPSISAPHRQA 292
 QY 326 YHLYETPEGVVSIGRSIYMTVPSTYASNI--LRPLSYAAADALSNFY---YPPVGAVTI 381
 DB 293 EESEFD-----AVIMAPLCQVKS--KIAKRGMPFLNTPEDVDPVLSVYIT 339
 QY 382 SYPOAIRDERLVDGLKGFQLHP---RTQGVETLGTIYSSSLFPNRAPKRVLILNY 438
 DB 340 TFKKESYKH-----PLESGFVLVSQDQKGLKRTLTSTSSMMPDRAPANNVLYITTFV 393
 QY 439 GGAAPPEI 446
 DB 394 GGSRRREL 401

RESULT 10
 D83800
 protoporphyrinogen IX and coproporphyrinogen III oxidase hemY [imported] - Bacillus h
 C:Species: Bacillus halodurans
 C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
 C:Accession: D83800
 R:Itakami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: A83650; M01D:2026314
 A:Accession: D83800
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-467 <STO>
 A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04923.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: hemY

Query Match 14.3%; Score 407.5; DB 2; Length 467;
 Best Local Similarity 28.2%; Pred No 2.4e-22;
 Matches 135; Conservative 90; Mismatches 210; Indels 43; Gaps 19;
 QY 76 IVGAGISGLCTIAQVMSANTPNLMTVEARDRAGNTTVERDGLTMEGPNST-OPS 130

```

Db      10  VIGGGMTGGLAAMALQAKRMPELEYELLEAAQAGGIGQTDVLEGFVERIPDSYLARK 69
Oy      131  DPMILTAAYVCGGLKDDLYVGGPNAPRFVLMKGLRVPESKL-----FDL-PFF--GLAMSTP 182
Db      70  ESMTRLARAREVLEGELVLRND-TGQAVITLKDQDLYVLPFGALIMGPIPEIGPIRSLRFSSTI 128
Oy      183  GKLRAGEFGPIGLRPPCPGHESEVGFVRNRLGCEVFELIPEPGSGVYVGDSPSKISMAAA 242
Db      129  GKLRAPADLVISFTIHENEVDLSLGFPRKRLRDEVDQLPEPLSLGIVKAGDDITLSLAT 188
Oy      243  FGRVYMLETGGSIIGTFPKALKERBSSTPKAPRDRRLRPRKGGTVSGFPRGIMLPDAIS 302
Db      189  FPGPQOLEKXISILRG-IRASGKOS-----QSQVKKRGKML-TPFRGGLOSTLEITE 240
Oy      303  ALRGLSLKLSMKLSSITKSEKG--YHLTETPEGVSSLASRSIVMTVPSVAASNILRPL 360
Db      241  QHLD--KAIKKKTEVRFCEEREKRRYRLTFN--DETQVYDH-VIITTPPYVTGRLPEPY 295
Oy      361  SYAAADALSNFYPPVGAVTISYPOAIRDEKRLVGLGELKGGQILHPRTQGVSTLTSTSS 420
Db      296  -----VDVYMREMKATTATATVAMAARFSEDL-HIPEYEGTFVPRKSRMYETACTMTH 347
Oy      421  SLFPNAPRAGVLLINTIGKAKNPEILSTKESLQLEVDVDRRLK-MLIKPKADQDLVGV 479
Db      348  KKKRHITTPGSHLVAVCYGKPGSESIYHETDAIVQYQKDQDQIMTVK---REPLFTKY 404
Oy      480  RVAPQAIQPELVGHLDTLSTAKAAMNDGDEBLFAGNYVSGVALGRCV-EGATVYAS 536
Db      405  TRKKKAMPYEYEGHAKTELVROELK-KAMPDIILAGGEGGIGLPRCCIDQGEAAVAT 461

```

```

Db 326 RDHML-----HGQVADESGIEEMKAITVHSKLIHGRBPSGHVLLRVFFKDLDPQKALA 378
QY 449 KTESQLEVVYDRLKRLMKIKPKRAQ-DPLVYGVWVPOAIPQFLVGLTSLSTAKAMND 507
Db 379 EAVHEVARLFG-----AQQDPLMHAYADMRGNKPNAYOVGHLDHIAIRALAPRN 427
QY 508 GLEGLEFLGNVYSGVALGRCVEGAYEVAEVTGFL 542
Db 428 ----IQVAGSSTYGVGIPDCVNAAGRTAARDVVAAL 458

RESULT 13
A56449
protoporphyrinogen oxidase (EC 1.3.3.4) - human
C:Species: Homo sapiens (man)
C:Date: 27-Oct-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
C:Accession: J04971; A56449; I38995
R:Puy, H.; Kobreau, A.M.; KosiPal, R.; Nordmann, Y.; Deybach, J.C.
Biochem. Biophys. Res. Commun. 226, 226-230, 1996
A:Title: Protoporphyrinogen oxidase: Complete genomic sequence and polymorphisms in the
A:Reference number: J04971; MUID:96400242
A:Accession: J04971
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-477 <PUY>
A:Cross-references: EMBL:X99450
R:Nishimura, K.; Taketani, S.; Inokuchi, H.
J. Biol. Chem. 270, 8076-8080, 1995
A:Title: Cloning of a human cDNA for protoporphyrinogen oxidase by complementation in vi
A:Reference number: A56449; MUID:95229621
A:Accession: A56449
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <NIS>
A:Cross-references: GB:D38537; NID:g854663; PIDN:BA07538.1; PID:g854664
R:Dalley, T.A.; Meissner, P.; Dailey, H.A.
J. Biol. Chem. 269, 813-815, 1994
A:Title: Expression of a cloned protoporphyrinogen oxidase.
A:Reference number: A49858; MUID:94117488
A:Accession: I38995
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-477 <RES>
A:Cross-references: EMBL:U26446; NID:g837327; PIDN:AAA7690.1; PID:g837328
C:Comment: This enzyme acts as the penultimate step in the heme biosynthetic pathway, at
C:Genetics:
A:Gene: GDB:PPOX; PPO
A:Cross-references: GDB:118852; OMIM:600923
A:Map position: 1q23-1q23
A:Introns: 29/3; 74/3; 157/3; 206/1; 269/3; 290/1; 329/3; 366/3; 416/3; 431/1
C:Superfamily: phytoene dehydrogenase
C:Keywords: oxidoreductase

Query Match 12.8%; Score 363.5; DB 2; Length 477;
Best Local Similarity 25.5%; Pred. No. 4.4e-19;
Matches 138; Conservative 83; Mismatches 170; Indels 151; Gaps 23;

QY 75 VIVGAGISGLCTAIOVMSANY-----PNIMVTEARDRAGCNITVE-RDGYLMEGPN 125
Db 6 VILGGISGL-----AASYHLSRAPCPKPKVYLVESSRLGWSVGPNGAPILPSPR 59
QY 126 SHOPSOPM---LTMAVDCGLKDDL--VLGDPNAP--RFLVMKGLRVPVSKITLDLPFFD 177
Db 60 GTRPGALGARLLLVSELGDLSEVLVPRGDHPAQNRFLLVGGALHLPGLRGL----- 115
QY 178 LMSIPGKLKLAGFPIGLRPSPP-----GHE--ESVDFVRNNGEVEF 218
Db 116 -----LRSPSPSKPLFMWAGLRELTKRKGEPDVTVSFAQRRLGPEVA 159
QY 219 ERLIEPFGCVVYVGPDSKLSKMAAFGKYWKLEETGGSIIGTFKAIKERSSTPKAPRDP 278

```

```

Db 160 SLADSLCRGVAFNAGNSRLSCLPSLFOAEQYHRSITLGLIGA-----GR 207
QY 279 LPRKGGTV-----GSFRKGLRMLPDALSARLGS-----LKLs----- 312
Db 208 TPQDSALIRKALAEKRSQMSLRGLLEMLPQALETHLTSGVSVLRQPYCGSLQAEGR 267
QY 313 WKLSITRSEKGVNHLVETPEGVVSLQSRISVYTVPSVYASNILRPLSYAADAISNFY 372
Db 268 MKV-SLRSS-----SLEADHVISAIYASVLSLELLPPEAAPLRLALSAT 310
QY 373 YPPGAVTISTPQAIIDERLVDGELKFGOLHPRIOGVETLTGITYSSLPFNR--APK 430
Db 311 AVSVAVNLQYOGANL-----PVQFGHLVPSSEDPVLGIYDVSAFPPQDSSPPG 362
QY 431 -RVLLNVYGGAKNPRLSKTESQLEVVYDRLKRLMKIKPKADPLV--VGVWVP---- 483
Db 363 LRVTVW-LGS-----SWLOTLEASGCVLSQELPQORADQAAATOLGLEKMSHCL 411
QY 484 ----QAIPOFLVGLHDLTLSTAKAAMNDNGLEGLFGNTVSGVALGRCVEGAYEVAE 538
Db 412 VHLKNCIPQYTTGLHWQKLSARQFLTAHRLPFLAGASY-EGAVVNDCTESGRQAAVS 470
QY 539 TG 540
Db 471 LG 472

RESULT 14
S68367
protoporphyrinogen oxidase (EC 1.3.3.4) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 19-May-2000
C:Accession: S68367; S65630
R:Dalley, T.A.; Dailey, H.A.; Meissner, P.; Prasad, A.R.K.
Arch. Biochem. Biophys. 324, 379-384, 1995
A:Title: Cloning, sequence, and expression of mouse protoporphyrinogen oxidase.
A:Reference number: S68367; MUID:96132666
A:Accession: S68367
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-477 <DAI>
A:Cross-references: EMBL:U25114; NID:g793939; PIDN:AAA96003.1; PID:g793940
R:Taketani, S.; Yoshinaga, T.; Furukawa, T.; Kohno, H.; Tokunaga, R.; Nishimura, K.;
Eur. J. Biochem. 230, 760-765, 1995
A:Title: Induction of terminal enzymes for heme biosynthesis during differentiation o
A:Reference number: S65629; MUID:95331315
A:Accession: S65630
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-14; 164-178 <TAK>
C:Genetics:
A:Genome: nuclear
C:Superfamily: phytoene dehydrogenase
C:Keywords: mitochondrion; oxidoreductase

Query Match 12.7%; Score 362.5; DB 2; Length 477;
Best Local Similarity 24.8%; Pred. No. 5.2e-19;
Matches 133; Conservative 76; Mismatches 188; Indels 139; Gaps 19;

QY 75 VIVGAGISGLCTAIO--VMSANYPNIMVTEARDRAGCNITVE-RDGYLMEGPNFSPSD 131
Db 6 IVLGGISGLAASHYLRGSPRPVILVESSKRLGWSIRISDSDAIFPLGFRGIRAPG 65
QY 132 PM---LTMAVDCGLKDDL--VLGDPNAP--RFLVMKGLRVPVSKITLDLPFFDLMSIPG 183
Db 66 ALGARLLLVSELGDLSEVLVPRGDHPAQNRFLLVGGTLPPLPSGLRGL----- 115
QY 184 KLRAGFPIGLRPSPP-----GHE--ESVDFVRNNGEVEF 224
Db 116 -----LRSPSPSKPLFMWAGLRELTKRKGEPDVTVSFAQRRLGPEVASLAMD 165

```

| | | | | |
|----|-----|-----------------|---|-----|
| QY | 423 | FPR----- | APKRGVLLVIGGANNPELSTKTSQLEVEVDNRKMLKPKADPL | 475 |
| | | : | : : : : : : | |
| Db | 326 | FARKKMGVLADEDP | LVLRFSVGRYDTEFLGRDADGLVNSRRIDLEAGRTAA--PV | 383 |
| | | : | : : : : : : | |
| QY | 476 | VGVGRVPAIQFLVGH | LDPLSTAKAAMDNGLEGLFGGNTVSGVALGRCVEGAEYVA | 535 |
| | | : | : : : : : : | |
| Db | 384 | ATERTKMDGLDQY | EVGH--HARVARVREHVAKPLGLVAGGAAIDSGVIPSISATAA | 441 |
| | | : | : : : : : : | |
| QY | 536 | SEVTGFL | 542 | |
| | | : | : | |
| Db | 442 | DQIRGDL | 448 | |
| | | : | : | |

CC This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
CC of the invention, which is tolerant to light-requiring herbicides. The
CC enzyme and its derivatives can be applied to construct plants with high
CC tolerance to light-requiring herbicides, useful in agriculture.
XX

SQ Sequence 548 AA:

Query Match 99.7%; Score 2836; DB 20; Length 548;
Best Local Similarity 99.8%; Pred. No. 8.7e-264;

Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MTTPIAHNPHTFHSSSPFLAINTSFIFPSSISKNSVNCNGMRTCSVAKDYTP 60
DB 1 mttcplianhpniftfhgssspflaintsfifssisknsvncngwrtcsvakdytp 60
QY 61 SSAVDGPPAALDCVYAGAGISGLCIAQVMSANYPMLMTTEARDRAGNTTVERDGYLW 120
DB 61 ssavdggppaaeldcvlyagagisglciaqymsanypmlmteardragnttverdgylw 120
QY 121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLDGPNAPRVLMKGLRPVPSKLTLDLPEFDLMS 180
DB 121 eegpnsfqpdsdpmlltmavdcgllkddlvldgpnaprflvkmglrppvskltdlpefdlms 180
QY 181 IPGKLRAAGFGPIGLRSPPGHEESVEQFVRNMGGEVEERLIEPFCSGYVGDPSKLSMK 240
DB 181 ipgklragfgaiglrpsppgheesveqfvrnmggeveerliepfcsgyvvgdpsklsmk 240
QY 241 AAGKWKLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKGQVWGSFRKGLRMLPDA 300
DB 241 aagkwkleeetgsiigtfkaikerstpkaprdlprpkgqvwsfrkglrmlpda 300
QY 301 ISARLGSKLMSKLSSTITKSEKGYHLTYETPEGVVSLQSRISVMTVPYSVANSILRPL 360
DB 301 isarlgsklmsklsstiksekgyhlyetpegvvsllqsrsvmtvpsvasnllrpl 360
QY 361 SVAADALSNFYPPYGVAVTISYPOAIRDRLVDELGFGQLHPRTOGVETLGTIYSS 420
DB 361 svaadalsnfyppvgavtlisypgealrderlvdelegfgqlhprtgvettlgtiys 420
QY 421 SLFPNRPKGRVLLNLTIGAKNPETLSTESQLEVEVDRDLKMLIKPKADPLVVGVR 480
DB 421 slfpnrpkgrvlllntyigaknpelstkesqlveevdrdlrkmllpkxadplvvgyr 480
QY 481 VMPQALPQFLVGHLDLTSTAKAAMNDNGLEGLFLGNTVYSGVALGRCEGAYEVAASEVTG 540
DB 481 vmpgalpqflvghldltstakaamndngleglflgntvysgvalgrcegayevasevtg 540
QY 541 FLSTRYAYK 548
DB 541 flstryayk 548
```

RESULT 2

AAW05201
ID AAW05201 standard; Protein: 548 AA.

AC AAW05201;

XX 16-JUN-1999 (first entry)

DE N. tabacum protoporphyrinogen oxidase mutant.

KW Protoporphyrinogen oxidase; light-requiring herbicide tolerance; enzyme.

OS Nicotiana tabacum.

XX WO9913087-A1.

XX 18-MAR-1999.

PF 10-SEP-1998; 98WO-JP04064.

XX

PR 11-SEP-1997; 97JP-0265084.

XX (NIPPY) NIPPON NOYAKU CO LTD.

PI Hirooka T, Horikoshi M, Mametsuka K;

DR WPI: 1999-215067/18.

DR N-PSDB: AAX28279.

PT protoporphyrinogen oxidase tolerant to light-requiring herbicides,

PS useful for constructing plants with such properties.

CC Example 5; Page 44-49; 56pp; Japanese.

CC This sequence is the Nicotiana tabacum protoporphyrinogen oxidase
CC of the invention, which is tolerant to light-requiring herbicides. The
CC enzyme and its derivatives can be applied to construct plants with high
CC tolerance to light-requiring herbicides, useful in agriculture.
XX

SQ Sequence 548 AA:

Query Match 99.7%; Score 2836; DB 20; Length 548;
Best Local Similarity 99.8%; Pred. No. 8.7e-264;

Matches 547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 MTTPIAHNPHTFHSSSPFLAINTSFIFPSSISKNSVNCNGMRTCSVAKDYTP 60
DB 1 mttcplianhpniftfhgssspflaintsfifssisknsvncngwrtcsvakdytp 60
QY 61 SSAVDGPPAALDCVYAGAGISGLCIAQVMSANYPMLMTTEARDRAGNTTVERDGYLW 120
DB 61 ssavdggppaaeldcvlyagagisglciaqymsanypmlmteardragnttverdgylw 120
QY 121 EEGPNSFQPSDPMLTMAVDCGLKDDLVLDGPNAPRVLMKGLRPVPSKLTLDLPEFDLMS 180
DB 121 eegpnsfqpdsdpmlltmavdcgllkddlvldgpnaprflvkmglrppvskltdlpefdlms 180
QY 181 IPGKLRAAGFGPIGLRSPPGHEESVEQFVRNMGGEVEERLIEPFCSGYVGDPSKLSMK 240
DB 181 ipgklragfgaiglrpsppgheesveqfvrnmggeveerliepfcsgyvvgdpsklsmk 240
QY 241 AAGKWKLEETGSGIIGTFKAIKERSSTPKAPRDLPRPKGQVWGSFRKGLRMLPDA 300
DB 241 aagkwkleeetgsiigtfkaikerstpkaprdlprpkgqvwsfrkglrmlpda 300
QY 301 ISARLGSKLMSKLSSTITKSEKGYHLTYETPEGVVSLQSRISVMTVPYSVANSILRPL 360
DB 301 isarlgsklmsklsstiksekgyhlyetpegvvsllqsrsvmtvpsvasnllrpl 360
QY 361 SVAADALSNFYPPYGVAVTISYPOAIRDRLVDELGFGQLHPRTOGVETLGTIYSS 420
DB 361 svaadalsnfyppvgavtlisypgealrderlvdelegfgqlhprtgvettlgtiys 420
QY 421 SLFPNRPKGRVLLNLTIGAKNPETLSTESQLEVEVDRDLKMLIKPKADPLVVGVR 480
DB 421 slfpnrpkgrvlllntyigaknpelstkesqlveevdrdlrkmllpkxadplvvgyr 480
QY 481 VMPQALPQFLVGHLDLTSTAKAAMNDNGLEGLFLGNTVYSGVALGRCEGAYEVAASEVTG 540
DB 481 vmpgalpqflvghldltstakaamndngleglflgntvysgvalgrcegayevasevtg 540
QY 541 FLSTRYAYK 548
DB 541 flstryayk 548
```

RESULT 3

AAW41608
ID AAW41608 standard; Protein: 539 AA.

AC AAW41608;

XX

| Query Match | 74.3%; Score 2113; DB 18; Length 539; |
|---------------------------|--|
| Best Local Similarity | 79.2%; Pred. No. 2.8e-194; |
| Matches 401; Conservative | 46; Mismatches 55; Indels 4; Gaps 1 |
| 47 | WRTGCSVANDYVPPSSANDVGGPAELIDCYIVAGISGLCIAOVMSANTY----PNTMYTTEA 102 |
| Db | 34 fklrsciaegpflssskldygessladvivgyslgclaqalackhndvasnvltvea 93 |
| Qy | 103 RRRAGGINTTVERDGLYMEGPNSSFOPSOPMTMAVDCGLKLDLVLGDPNAPREYVLMKKK 162 |
| Db | 94 rdtvggnlttverdyglweegpnsfjpspblltmavdsglkdldlvlgdpnaprfvfwegk 153 |
| Qy | 163 LRVPESKLTLDLPEDIMSLIPGKIRAGFGPIGLRPSFGHEBSVEQVRRNRLGGEVERLI 222 |
| Db | 154 lrpypskpctdlpfdlmslagnklraqfgalgltrpppyveesveefvrnlgaevefeti 213 |
| Qy | 223 EPPGSGVYVGDSDSKSMKAAFGVWMLLEETGSGIIGGTFAIKERSSSTPKAPDPLPKP 282 |
| Db | 214 efcfsgvyvagdpskismkaafgrvwlleeigsgiiyggtfltiqernktbkpprdpdlpnp 273 |
| Qy | 283 KGVTSVSFRKGLMLDPAISARLGSKLKLSWKLSSITKSEKGYHLVYETTPGEGVSLSOR 342 |

| | | | |
|-----------------------|----------|---|-----|
| Db | 274 | kqgtvsgtrfkltlmlpeaianslgsnvlkslkslcklignngynlftcepegmwslqsr | 333 |
| OY | 343 | SIYMTVPYVYASNTILRPLSLVAADALSNFYPPYVGAVTISTPOEALIDERLVDGELKGF | 402 |
| Db | 334 | svmtlphvsaenlllhpisaaadalsqfypypvasvtsypkcalrkeclldgelkgif | 393 |
| OY | 403 | OLHPTQGVLELGTITSSLEPNNAPKGRVLLNLTIGAKNPDLSTKSTESOLVEVVDRL | 462 |
| Db | 394 | qlhprsggletlgtlyssllpmtapsgrvlllthysgaatntgllsktegeiveavrdl | 453 |
| OY | 523 | ALGRCEGATYEYASFEVTFGLSRAYK | 548 |
| Db | 514 | algrcegayeavaevkeflsqayk | 539 |
| RESULT 4 | | | |
| AAW41609 | ID | AAW41609 standard; Protein: 539 AA. | |
| XX | AC | AAW41609; | |
| XX | DT | 20-APR-1998 (first entry) | |
| XX | DE | Cotton protox-1. | |
| XX | KW | Protoporphyrinogen oxidase-1; protox-1; promoter; cotton; | |
| XX | KM | herbicide resistance; breeding programme; probe; gene isolation; | |
| XX | OS | genomic mapping. | |
| XX | PN | Gossypium hirsutum. | |
| XX | PD | W09732028-A1. | |
| XX | PF | 04-SEP-1997. | |
| XX | PR | 27-FEB-1997; 97WO-US03343. | |
| XX | PR | 21-JUN-1996; 96US-0020003. | |
| XX | PR | 28-FEB-1996; 96US-0012705. | |
| XX | PA | 28-FEB-1996; 96US-0013612. | |
| XX | PI | (NOVS) NOVARTIS AG. | |
| XX | PI | Johnson MA, Volrath SL, Ward ER; | |
| XX | DR | WPI. 1997-489209/45. | |
| XX | DR | N-PSDB; AAW04313. | |
| XX | PT | DNA containing a plant proto-porphyrinogen oxidase gene promoter - | |
| XX | PT | optionally linked to a heterologous gene, especially to express | |
| XX | PS | herbicide-resistant enzymes, and plants containing such constructs | |
| XX | PS | Claim 35; Pages 80-82; 114pp; English. | |
| XX | CC | The present sequence is cotton protoporphyrinogen oxidase-1 | |
| XX | CC | (protox-1). | |
| XX | CC | The protox-1 promoter can be used to express herbicide resistant | |
| XX | CC | enzymes, specifically protox, i.e. a plant tissue, plant or progeny | |
| XX | CC | containing a chimeric gene of the promoter and a heterologous | |
| XX | CC | coding sequence. The plant can also be used in breeding programmes. | |
| XX | CC | Also hybridising fragments of the protox coding sequence can be | |
| XX | CC | used as probes, e.g. to isolate related genes or for genomic | |
| XX | CC | mapping. | |
| XX | Sequence | 539 AA; | |
| Query Match | 74.3%; | Score 2113; DB 18; Length 539; | |
| Best Local Similarity | 79.2%; | Pred. No. 2.8e-194; | |

| Matches | 401; Conservative | 46; Mismatches | 55; Indels | 4; Gaps | 1; |
|----------|---------------------|--|------------|---------|----|
| QY | 47 | WRTRCVAKDYTVSSAVDGPAAELDCVIVGAGISGLCTAQMVSANY---- | PNLMWTEA | 102 | |
| Db | 34 | fklrcslaegptlssskldgessiadcvlvgglsqcltaqalackhdvasnvlvtea | 93 | | |
| QY | 103 | RDRAGGNTTVERDGYLMEBPNFQSDPMLTMADCGLKDLYLVGDPMARFVLMKRG | 162 | | |
| Db | 94 | rdtvggnlttverdyjlweegpnsfqpdpbltmavdsglkddlvigdpnaprfvlwiegk | 153 | | |
| QY | 163 | LRPVPSKLTDLPEFDLMSIPGKIRACFGPIGLRPPGHESEVDFVRNLDGEVERLI | 222 | | |
| Db | 154 | lrpvpksptldlptfdlmslraglragfgalirpppyveesveefvrnlgaeverfi | 213 | | |
| QY | 223 | EPFGSGVYVGDPSKLSMKAFFGVWKLLEETGSIIGCTFKAIKERSSTPKARDPRLPKP | 282 | | |
| Db | 214 | epfcsgyvagdpsklsmkafgrvwkleeigsligqcltqerlnktpkprdpripxp | 273 | | |
| QY | 283 | KGQTVGSFRRGLRMLDPAISARLSKLSWKLSITKSEKGYHLVETPEGVLSQSR | 342 | | |
| Db | 274 | kgqtvgsfrrglmlpealanslgnvklswklstklngngynltfclpegmwslqsr | 333 | | |
| QY | 343 | SIVMTVPVYASNILRPLSVAAADALSNFYPPVCAVITSYFOEAIKRDRLVDGELKFG | 402 | | |
| Db | 334 | svmtlpsvhasnllhplsaadaalsqfyyppvasvltvypkealirkeclldgelkfg | 393 | | |
| QY | 403 | QLHPTQGVETLCTIYSSSLFPNRAKGRVLLNTYGAKNPBILSKTESQLEVEVDRDL | 462 | | |
| Db | 394 | qlhprsqetlctiyssslfnpnrapsgvlllnlygagntngllsktegeelveavddl | 453 | | |
| QY | 463 | RKMLIRPKADPLVGVWRVWPAQIIPQFLVGHLDLTSTAKAANDNGLEGLFGNYYVSGV | 522 | | |
| Db | 454 | rkmlirpnakdpilvgvwrwpkaiipqflvghldlldsakmalrdsqfnglflgnyvsgv | 513 | | |
| QY | 523 | ALGRCEYAEVYASEVTGFLSRYAYK | 548 | | |
| Db | 514 | algrceyaeveaevkeflsqayk | 539 | | |
| RESULT | 5 | | | | |
| AAW25740 | | | | | |
| ID | AAW25740 | standard; Protein; 539 AA. | | | |
| XX | AAW25740; | | | | |
| AC | | | | | |
| XX | | | | | |
| DT | 01-MAR-1998 | (first entry) | | | |
| XX | | | | | |
| DE | | Cotton protoporphyrinogen oxidase (protox-1). | | | |
| XX | | | | | |
| KM | | Protox-1; protoporphyrinogen oxidase; inhibitor; cotton; | | | |
| XX | | herbicide tolerance; herbicide resistance; transgenic plant. | | | |
| OS | | Gossypium hirsutum L. | | | |
| XX | | | | | |
| XX | | | | | |
| FH | Key | Location/Qualifiers | | | |
| FT | Misc-difference 365 | /note- "substitution of Pro-365 by another amino | | | |
| FT | | acid, especially Ser, provides a modified | | | |
| FT | | protox tolerant to a herbicide (Claims | | | |
| FT | | 51-52)" | | | |
| FT | Misc-difference 428 | | | | |
| FT | | /note- "substitution of Tyr-428 by another amino | | | |
| FT | | acid, especially Cys or Arg, provides a | | | |
| FT | | modified protox tolerant to a herbicide | | | |
| FT | | (Claims 53-54)" | | | |
| XX | | | | | |
| XX | | | | | |
| PN | WO9732011-A1. | | | | |
| XX | | | | | |
| PD | 04-SEP-1997. | | | | |
| XX | | | | | |
| PE | 27-FEB-1997; | 97WO-US03313. | | | |
| XX | | | | | |
| PR | * 21-JUN-1996; | 96US-0020003. | | | |

| | | | | | |
|----|---|--|----------|-----|--|
| PR | 28-FEB-1996; | 96US-0012705. | | | |
| PR | 28-FEB-1996; | 96US-0013612. | | | |
| XX | | | | | |
| PA | (NOVS) NOVARTIS AG. | | | | |
| XX | | | | | |
| PI | Heifetz PB, Johnson MA, Potter SL, Volrath SL, Ward ER; | | | | |
| XX | | | | | |
| DR | WPI: 1997-448683/41. | | | | |
| XX | N-PSDB; AAT86123. | | | | |
| PT | New DNA encoding plant protoporphyrinogen oxidase enzyme - and | | | | |
| PT | herbicide resistant mutants, useful to prepare plants resistant to | | | | |
| PT | herbicide which therefore kills undesired vegetation only | | | | |
| XX | | | | | |
| XX | Claim 7; Page 139-141; 196pp; English. | | | | |
| XX | | | | | |
| CC | This protein comprises cotton protoporphyrinogen oxidase | | | | |
| CC | (protox-1), an enzyme that catalyses the oxidation of | | | | |
| CC | protoporphyrinogen IX to protoporphyrin IX. Its amino acid | | | | |
| CC | sequence was deduced from an cotededon Protox-1 cDNA clone (see | | | | |
| CC | AAT86123). Sites within the cotton Protox-1 gene have been | | | | |
| CC | identified that can be mutated to encode a modified protox that is | | | | |
| CC | resistant to protox inhibitors and hence tolerant of certain | | | | |
| CC | herbicides. Plants, especially crop plants, may be engineered for | | | | |
| CC | resistance to protox inhibitors via mutation of the native protox | | | | |
| CC | gene to a resistant form, or they may be transformed with a gene | | | | |
| CC | encoding an inhibitor-resistant form of a plant protox enzyme, such | | | | |
| CC | as claimed forms from wheat, soybean, cotton, sugarcane, oilseed | | | | |
| CC | rape, rice and sorghum (see AAW25738-48). Application of herbicide | | | | |
| CC | will then kill undesired vegetation only. Protox enzymes can also | | | | |
| CC | be expressed in transformed host cells and used to identify | | | | |
| CC | inhibitors of protox enzyme activity, i.e. herbicide candidates, or | | | | |
| CC | to design herbicide tolerant forms of the enzyme. | | | | |
| XX | | | | | |
| SQ | Sequence | 539 AA; | | | |
| | Query Match | 74.3%; Score 2113; DB 18; Length 539; | | | |
| | Best Local Similarity | 79.2%; Pred. No. 2, 8e-194; | | | |
| | Matches | 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1; | | | |
| QY | 47 | WRTRCVAKDYTVSSAVDGPAAELDCVIVGAGISGLCTAQMVSANY---- | PNLMWTEA | 102 | |
| Db | 34 | fklrcslaegptlssskldgessiadcvlvgglsqcltaqalackhdvasnvlvtea | 93 | | |
| QY | 103 | RDRAGGNTTVERDGYLMEBPNFQSDPMLTMADCGLKDLYLVGDPMARFVLMKRG | 162 | | |
| Db | 94 | rdtvggnlttverdyjlweegpnsfqpdpbltmavdsglkddlvigdpnaprfvlwiegk | 153 | | |
| QY | 163 | LRPVPSKLTDLPEFDLMSIPGKIRACFGPIGLRPPGHESEVDFVRNLDGEVERLI | 222 | | |
| Db | 154 | lrpvpksptldlptfdlmslraglragfgalirpppyveesveefvrnlgaeverfi | 213 | | |
| QY | 223 | EPFGSGVYVGDPSKLSMKAFFGVWKLLEETGSIIGCTFKAIKERSSTPKARDPRLPKP | 282 | | |
| Db | 214 | epfcsgyvagdpsklsmkafgrvwkleeigsligqcltqerlnktpkprdpripxp | 273 | | |
| QY | 283 | KGQTVGSFRRGLRMLDPAISARLSKLSWKLSITKSEKGYHLVETPEGVLSQSR | 342 | | |
| Db | 274 | kgqtvgsfrrglmlpealanslgnvklswklstklngngynltfclpegmwslqsr | 333 | | |
| QY | 343 | SIVMTVPVYASNILRPLSVAAADALSNFYPPVCAVITSYFOEAIKRDRLVDGELKFG | 402 | | |
| Db | 334 | svmtlpsvhasnllhplsaadaalsqfyyppvasvltvypkealirkeclldgelkfg | 393 | | |
| QY | 403 | QLHPTQGVETLCTIYSSSLFPNRAKGRVLLNTYGAKNPBILSKTESQLEVEVDRDL | 462 | | |
| Db | 394 | qlhprsqetlctiyssslfnpnrapsgvlllnlygagntngllsktegeelveavddl | 453 | | |
| QY | 463 | RKMLIRPKADPLVGVWRVWPAQIIPQFLVGHLDLTSTAKAANDNGLEGLFGNYYVSGV | 522 | | |
| Db | 454 | rkmlirpnakdpilvgvwrwpkaiipqflvghldlldsakmalrdsqfnglflgnyvsgv | 513 | | |

OY 523 ALGRCEGAYEVAEVTGELSRVAYK 548
 DB 514 algrcegayevaevkeflsqayk 539

RESULT 6

AA872907 standard; Protein: 539 AA.

AC AAB72907;

DT 16-MAY-2001 (first entry)

DE Cotton protoporphyrinogen oxidase SEQ ID NO: 16.

KM Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;

OS Gossypium hirsutum.

XX WO200112825-A1.

XX 22-FEB-2001.

PF 30-JUN-2000; 2000WO-EP06127.

PR 13-AUG-1999; 99US-0373691.

PA (SYNG-) SYNGENTA PARTICIPATIONS AG.

PI Johnson MA, Volrath SL, Helfetz PB, Law MD;

DR WPI: 2001-234914/24.

XX N-PSDB: AAF6579.

PT Plant DNA molecules encoding herbicide-tolerant forms of

PS protoporphyrinogen oxidase which are useful for rationally designing

XX transgenic plants and seeds -

XX Claim 22; Page 184-187; 228pp; English.

CC The present invention provides the protein and coding sequences of a

CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,

CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen

CC oxidase (protox) enzyme. Examples of these mutants are shown in

CC AAB72920-AAB72926. They are useful as they enable the production of

CC herbicide-tolerant plants and seeds. The present sequence is a protox

CC protein.

XX Sequence 539 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

OY 283 KQTVGSEFRKGLRMLPDALISRLGSKLSMKLSSITSEKGYHLTYETPEGVSLQSR 342
 DB 274 kgqtvgsfrkglrmlpdalislrgsklsmsklsstlsgngyhltyetpvgvslqsr 333
 OY 343 SLVMTVPYSVANSNIIIRPLSVAAADALSNFYPPVAVTISYPOAIRDERLVJGELKFG 402
 DB 334 svmtlpsvasnllhplsaaadalsqfypvasvcsypkkaierkeclldgelkfg 393
 OY 403 QLRPRGCVETLGTIVSSSLFPNRAKGRVLLNTYGAKNPETLSTESQIVENVDRDL 462
 DB 394 qlrpsrgletgtivssslfpmrapsgrvllnyggatctgllskdegleveavdrdl 453
 OY 463 RKMLIKRAQDELVGVVWVWQAIPOFLVGHLDPLSTAKAAMNDGELGLFGNTVSGV 522
 DB 454 rkmlinrakdpdvlygvvwpkaiqpflvghldldskamalrdsqfllfgnyvsgv 513
 OY 523 ALGRCEGAYEVAEVTGELSRVAYK 548
 DB 514 algrcegayevaevkeflsqayk 539

RESULT 7

AA872921 standard; Protein: 539 AA.

AC AAB72921;

DT 16-MAY-2001 (first entry)

DE Cotton protoporphyrinogen oxidase Y428H mutant.

KM Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;

XX soybean; sugar beet; oilseed rape; sugar cane; mutant; mutain.

OS Gossypium hirsutum.

XX Synthetic.

XX Key Location/Qualifiers

XX Msc-difference 428

XX /note="Wild-type Tyr substituted by His"

XX WO200112825-A1.

XX 22-FEB-2001.

XX 30-JUN-2000; 2000WO-EP06127.

XX 13-AUG-1999; 99US-0373691.

XX (SYNG-) SYNGENTA PARTICIPATIONS AG.

XX Johnson MA, Volrath SL, Helfetz PB, Law MD;

XX WPI: 2001-234914/24.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 74.3% Score 2113; DB 22; Length 539;
 Best Local Similarity 79.2%; Pred. No. 2.8e-194;
 Matches 401; Conservative 46; Mismatches 55; Indels 4; Gaps 1;
 OY 47 WTRRCVAKDYVPSAAVDGPAALDCVYAGISGLCIAQVMSKNY----PLNMTVEA 102
 DB 34 fkrtrcslaeapltsskldgsssladcvlyvgglsqldlaqalatkhrdvasvilylea 93
 OY 103 RDRAGNITTVREDGYLWEEGPNSEFOPSDPMLTMADVDCGLKIDLVLGDPNAPREVLMK 162
 DB 94 rdrvagnitvtredgylwegpnstqpsdpmltmavdsqldkdlvlgdpnprfvlwsgk 153
 OY 163 LRFPVSKLDLPFPDLMSPGKIRAGFGPGLRPSPPGHEEVEQFVRNNGEVEFELI 222
 DB 154 lrpvsksldlpfdlmslsgkrlraglgaqlrpppyeeveefvtrnlgavevtefl 213
 OY 223 EPRCSGVVYVGDPSKLSMKAARFGVWKLSEETGSGIIGTEFKAIKESSTPKAPRDPRLPK 282
 DB 214 eprcsgvvyvgdpskslsmkaarfgvwlseegtsigltfkigernkpkpdpdpk 273

Claim 22; Page 7; 228pp; English.
 The present invention provides the protein and coding sequences of a
 number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
 sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
 oxidase (protox) enzyme. Examples of these mutants are shown in
 AAB72920-AAB72926. They are useful as they enable the production of
 herbicide-tolerant plants and seeds. The present sequence is a mutant
 protox protein.
 Note: The present sequence is not shown in the specification but is
 derived from that given in SEQ ID NO: 16 (see AAB72907).

Sequence 539 AA;

[illegible]

| | | |
|---|---|--|
| PI | Johnson MA, Voltrath SL, Helfetz PB, Law MD: | |
| DR | WPI: 2001-234914/24. | |
| XX | | |
| PT | plant DNA molecules encoding herbicide-tolerant forms of | |
| PT | protoporphyrinogen oxidase which are useful for rationally designing | |
| PT | new inhibitory herbicides and for producing herbicide-tolerant | |
| PT | transgenic plants and seeds - | |
| PS | Claim 22: Page -: 228pp; English. | |
| XX | | |
| CC | The present invention provides the protein and coding sequences of a | |
| CC | number of herbicide-tolerant forms of wheat, soybean, rice, sorghum, | |
| CC | sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen | |
| CC | oxidase (protox) enzyme. Examples of these mutants are shown in | |
| CC | AA872920-AA872926. They are useful as they enable the production of | |
| CC | herbicide-tolerant plants and seeds. The present sequence is a mutant | |
| CC | protox protein. | |
| CC | Note: The present sequence is not shown in the specification but is | |
| CC | derived from that given in SEQ ID NO: 16 (see AA872907). | |
| XX | | |
| XX | Sequence 539 AA: | |
| SO | | |
| Query Match | 74.0%; Score 2104; DB 22: Length 539; | |
| Best Local Similarity | 79.1%; Pred. No. 2.1e-193; | |
| Matches 400; Conservative 46; Mismatches 56; Indels 4; Gaps | | |
| OY | 47 MRRRCVADYDVSSAVSDGPAELDCTYVAGTSGICIAQVMSNY----PMLATPEA 102 | |
| DB | 34 FKIRCSLAEGPIASLKKIDGESSLDCVLYGSGICLQGLACKHDAVSIVTLEA 93 | |
| OY | 103 RDAAGNITTVPRGGYLMWEGSPSPQSDPMITWAVCGIKDDLYGDPNAPRFLWKGK 162 | |
| DB | 94 RLVYGNILTVTERDGYLWEEGNSFSPSPDLITMAVDSGLKDDLVLGDPNPRFLWEGK 153 | |
| OY | 163 LRVYPSKLLDLPRFLMSIPGKIRAEFGYIGLRSPFGHESEYQVVRNKLGGYFERLI 222 | |
| DB | 154 LRVYPSKPLDLPRFLMSIAGKLRAGIGALYIRPPPYEVEELVRINLGAETFLI 213 | |
| OY | 223 EPPGSGVYVGDPSKISKKAAPGVYMKLEFTGSGIIGTFPAIKENSGTAKAPDRPLPKP 282 | |
| DB | 214 EPCISGVYVAGDPSKISKKAAGFYVMVLEEISGSLIGLFLIETKIRKCPKPRDRLPKP 273 | |
| OY | 283 KGGTVGSFPFKGLRMLPDAISARLGSKLKLSMKLSTITSKGGYHLVETPBGVYSLDSR 342 | |
| DB | 274 KGLTVGSIFTKGLTLPALANSGYVVKLSWLSLCKLNGSYNLTLFTEPGMWSYQSR 333 | |
| OY | 343 SIWYVYSVAYNILRLSVAAADALSNFYEPVCAVTIYSPQEAIRDLRLVDGELKFGC 402 | |
| DB | 334 SYWVTLSPVHSNLIHLPISAAADAISGLFYVPVAASYVPEAKRTKCLIDGELKFG 393 | |
| OY | 403 QLIHRTQGETLGLTVSSSLFPNRAIKGRVLLNLVYIGAKNPILSKTESQLEVEDRL 462 | |
| DB | 394 QLPHPRGELTGLTVSSSLFPNRAISGVVLLNLVYIGALNTGLSKTEGELVEAVDRDL 453 | |
| OY | 463 RKRLIKPKADPIVYGRVWQALIPQFLCGLHDTISYAKKAMNDLDELFLFGNGVYSGV 522 | |
| DB | 454 TKMLHPNAKDPVLVGYRVKPAKPIGLVGHDLIDKAKMALDSGFLNGLYGNYSGV 513 | |
| OY | 523 ALGRCEGAYEYASDYVGFISRAVX 548 | |
| DB | 514 ALGRCEGAYEYAAEVKELISQAYX 539 | |
| RESULT | 9 | |
| AA872923 | standard; Protein: 537 AA. | |
| XX | AA872923; | |
| XX | | |
| XX | 16-MAY-2001 (first entry) | |
| XX | | |

```

DE Arabidopsis protoporphyrinogen oxidase A2201 mutant.
XX
XX Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutlein.
XX
XX Arabidopsis thaliana.
OS
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 220 /note="wild-type Ala substituted by Ile"
FT
XX
XX W0200112825-A1.
XX
XX 22-FEB-2001.
XX
XX 30-JUN-2000; 2000MO-EP06127.
XX
XX 13-AUG-1999; 99US-0373691.
XX
XX (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
XX Johnson MA, Volrath SL, Helfetz PB, Law MD;
PI WPI; 2001-234914/24.
XX
XX Plant DNA molecules encoding herbicide-tolerant forms of
PT protoporphyrinogen oxidase which are useful for rationally designing
PT new inhibitory herbicides and for producing herbicide-tolerant
PT transgenic plants and seeds -
XX
XX Claim 23; Page -: 228pp; English.
XX
XX The present invention provides the protein and coding sequences of a
CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC oxidase (protox) enzyme. Examples of these mutants are shown in
CC AAB72920/AAB72926. They are useful as they enable the production of
CC herbicide-tolerant plants and seeds. The present sequence is a mutant
CC protox protein.
CC Note: The present sequence is not shown in the specification but is
CC derived from that given in SEQ ID NO: 2 (see AAB72901).
XX
XX
XX Sequence 537 AA;
SQ
Query Match 73.1%; Score 2078; DB 22; Length 537;
Best Local Similarity 75.4%; Pred. No. 6.5e-191;
Matches 404; Conservative 50; Mismatches 72; Indels 10; Gaps 4;
OY 22 LAFLNRISFTIPSSISRSNVCNCGW---RTRCSYAKDYTPSSAVDGPAAEL--DCVI 76
DB 3 lslrlptqslpsfksfn-llrnykplrllcsyagpvtgsskleggggtlttdevl 61
OY 77 VGAGISGICIAQVMSANY----PNLWTEARDRAGNTTVERDGYLMEGSPNSFPQSDP 132
DB 62 vggsgisgiclaqalatkhpdaapnlvteakrvgnlltteengflwsegpsfqspsp 121
OY 133 MTTMNVDCGLDLDVLDPNARREVLMKGLRPVPSKTLDPFDLMSIFGKLRAGFRI 192
DB 122 mltmnvdcglldvldpnaarrevlmkglrpvpskltldpfdlmsifgklragfria 181
OY 193 GLRSPRGHESVBOFVRNMLGGEVERLIEPFCSGVYVGDPSKLSMKAAGKVMLEET 252
DB 182 glrsprghesvbofvrnmlggevrliepfcsgyvlgdpsklsmkaafigkvmleeg 241
OY 253 GGSITIGGFFKAKRKSSTPRKAPRDLPRKQGTGSPFRKRLPDAISARIGSKLKS 312
DB 242 ggsitiggtfkaikrksstprkaprldprkqgtvgstfrkrlpdaisarigsklks 301
OY 313 WTLSSITSEKGGVHLYTERPBGVSLOSRSIVMVPYVSNTLRBLSAADAASNY 372
DB 302 wtlssitsekggvhylyterpbgvslsrsivmvpysntlrblsaadaasny 361

```

```

OY 373 YPPVCAVTISYPOEATRDERLYDGEIKGFQGLHPRTOGVETICTIYSSSLFPNRPKGRV 432
DB 362 yppvaavtsiypokeatrltcolldgellkgfqlhprtgvetlgltyssslfpnrappri 421
OY 433 LLNMYIGANKRPETLSKTESOLVEVVDRLKRMILKPRADDPVGVVRWPAIPQFLVG 492
DB 422 llmnyigstntglleskegelveavdrldlrmilkpnsldpklqvrwvpqalpqflvg 481
OY 493 HLDLTSTAKAAMMDNGLEGFEGNYSVGVALGRCVGEAGVAYEVTEVGLSYAVK 548
DB 482 hldltstaksslesgyegllfgnysvagalgrcvgeagayetalerlnmsryek 537
RESULn 10
AAR90295
ID AAR90295 standard; Protein; 537 AA.
XX
XX AAR90295;
AC
XX
XX 02-JUL-1996 (first entry)
DE
XX
XX Protein having protoporphyrinogen oxidase activity.
XX
XX Arabidopsis sp.
XX
XX Variegate porphyria.
XX
XX Arabidopsis sp.
XX
XX W09534659-A1.
XX
XX 21-DEC-1995.
XX
XX 08-JUN-1995; 95MO-IB00452.
XX
XX 16-JUN-1994; 94US-0261198.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX Volrath S, Ward ER;
PI WPI; 1996-049687/05.
XX
XX N-PSDB; AAT11673.
XX
XX DNA encoding eukaryotic proto:porphyrinogen oxidase and herbicide
PT resistant mutants - used to make herbicide resistant plants and for
PT diagnosis and treatment of variegate porphyria
XX
XX Claim 6; Page 71-73; 118pp; English.
XX
XX DNA encoding a modified protoporphyrinogen oxidase (PPO) or a
CC chimeric gene comprising a promoter (pref. active in a plant) linked
CC to such a DNA or DNA encoding a wild type PPO can be used (1) to
CC impart herbicide resistance to plants; (2) for treating and
CC diagnosing deficient PPO activity in animals (esp. variegate
CC porphyria); and (3) for the production of recombinant PPO which is
CC useful as an assay reagent and in rational design of new inhibitory
CC herbicides. Herbicide resistant PPO genes can also be used to
CC select plants transformed with a transgene and probes derived from
CC the genes can be used to quantify levels of PPO mRNA.
XX
XX
XX Sequence 537 AA;
SQ
Query Match 73.0%; Score 2075; DB 17; Length 537;
Best Local Similarity 75.4%; Pred. No. 1.3e-190;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;
OY 22 LAFLNRISFTIPSSISRSNVCNCGW---RTRCSYAKDYTPSSAVDGPAAEL--DCVI 76
DB 3 lslrlptqslpsfksfn-llrnykplrllcsyagpvtgsskleggggtlttdevl 61
OY 77 VGAGISGICIAQVMSANY----PNLWTEARDRAGNTTVERDGYLMEGSPNSFPQSDP 132

```


PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161094.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 73.0%; Score 2075; DB 21; Length 537;
Best Local Similarity 75.4%; Pred. No. 1.3e-190;
Matches 404; Conservative 49; Mismatches 73; Indels 10; Gaps 4;

QY 22 LAFLNRTSFIPESSISKRNVCNCGW---RTKRSVAKDITYVSSANDGSPADEL--DCYI 76
DB 3 LSLrlptqslpfskpn-lrlnvkplrlfrcsvagpplvssksklegggattltcdvyl 61
QY 77 VGAGISGICIAQVMSANY---PNLWTEARDRAGCNITTVREDGYLMEGPNFQPSDP 132
DB 62 vggisgiclaqalackhpdapnllvteaktvsgnlltreangflweegpnstfgpsp 121
QY 133 MLTMAVDCGLKDDLVLGDPAAPRFVLMGKLRPVPSKLTDLPEFDIMSTPGKIRAGFGPI 192
DB 122 mltnvvdsglkddlvlgdplaprfvlwngklrpypskltdlpefdlmslgyktragfgal 181
QY 193 GLRSPRGHEESVDFVRRNLGSEFERLIEPFGSVYVGDPSKLSMKAPKVMKLEET 252
DB 182 glrpsrgheesveefvrrnlgdevferllieptcsvysgdpsksmkaafigvwxlegn 241
QY 253 GCSILIGTFKAIKERSTPKAPRDLPRPKGQTVGSFRKGLMLPDATASRLGSKLKS 312
DB 242 ggsiligtfaikergnnpaerdprrlpkpgqtvsfrfkglmlpeasrlsgkvks 301
QY 313 WKLSTKSEKGYHLTYETPEGVSLQSRIVMTVPSTVANSILRPLSVAAADALSNTY 372
DB 302 wklsigtlesggyntltyetpdlvsavgskvmtvpslvasgllrplseasanaalskly 361
QY 373 YPPVGNATISYPOEARDELDVNDGELKGFQOLHPRFGVETTGITYSSLPFRAPKGV 432
DB 362 yppvaavsisypkearlteclldgelkfygqlhprtgvecltyssslfprappgrrl 421
QY 433 LLNLYIGAKNPILSKTESQLVEVVDRLRKMLIRPKADPLVWGVVRWPAIPQFLVG 492
DB 422 lllylgsstntglslkssegeivaeavdrdrdkmlkpnstdpklgyrrwpaipqflvg 481
QY 493 HLDLTSTAKAAMNDGLELFTLGNNVSGVALGRVCGAVEVASEVTGFLSRYAYK 548
DB 482 hfdlldtakselssgyegflfgnyvaagvalgrcvegeyetalevnlnmsryayk 537
RESULT 15
AAB72901
ID AAB72901 standard; Protein; 537 AA.
XX
AC AAB72901;
XX
DT 16-MAY-2001 (first entry)
XX
XX Arabidopsis protoporphyrinogen oxidase SEQ ID NO: 2.
DE Arabidopsis protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
XX
KM soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein.
XX
OS Arabidopsis thaliana.
XX
PN WO200112825-A1.
XX
PD 22-FEB-2001.
XX

| | |
|----------|---|
| PF | 30-JUN-2000; 2000WO-EP06127. |
| XX | |
| PR | 13-AUG-1999; 99US-0373691. |
| XX | |
| PA | (SYNG-) SYNGENTA PARTICIPATIONS AG. |
| XX | |
| PI | Johnson MA, Voltrath SL, Helfetz PB, Law MD; |
| XX | |
| DR | WPI; 2001-234914/24. |
| XX | |
| N-PSDB; | AAE76571. |
| XX | |
| PT | Plant DNA molecules encoding herbicide-tolerant forms of |
| PT | protoporphyrinogen oxidase which are useful for rationally designing |
| PT | new inhibitory herbicides and for producing herbicide-tolerant |
| XX | |
| XX | transgenic plants and seeds - |
| PS | |
| XX | Claim 23; Page 137-140; 228bp; English. |
| XX | |
| CC | The present invention provides the protein and coding sequences of a |
| CC | number of herbicide-tolerant forms of wheat, soybean, rice, sorghum, |
| CC | sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen |
| CC | oxidase (protox) enzyme. Examples of these mutants are shown in |
| CC | AAE72920-AAE72926. They are useful as they enable the production of |
| CC | herbicide-tolerant plants and seeds. The present sequence is a protox |
| CC | protein. |
| XX | |
| XX | |
| Sequence | 537 AA; |
| 50 | |

OM of: US-09-508-418-2 to: GenBank.* out_format : pfs

Date: Jul 3, 2001 11:36 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framed_p2n.model -DEV=xlp
-O=/cgn2_1/USPRO.spool/US09508418/runcat_03072001.100017.24438/app-query.fasta_1.613
-DB=GenBank -OEM=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -ICOPT=0.000 -ICOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=bloms62 -TRANS=human4.0.cd1
-LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09508418.@CGN1_1.5606 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPRY -WAIT -THREADS=1

Search information block:

Query: US-09-508-418-2

Query length: 548

Database: GenBank.*

Database sequences: 1344157

Database length: -856060004

Search time (sec): 1837.380000

Score list:

| Sequence | Strd Orig | ZScore | EScore | len | Documentation |
|------------------|-----------|---------|----------|--------|----------------------------------|
| gb_p11:AF044128 | + 2832.00 | 3442.64 | 1.6e-183 | 1874 | AF044128 Nicotiana tabacum prc |
| gb_p11:AB020501 | + 2824.00 | 3434.03 | 4.8e-183 | 1647 | AB020501 Nicotiana tabacum mrn |
| gb_p13:NTY13465 | + 2821.00 | 3429.16 | 9.0e-183 | 1892 | Y13465 Nicotiana tabacum mrn |
| gb_p14:STA225107 | + 2565.50 | 3117.46 | 2.1e-165 | 1976 | AJ225107 Solanum tuberosum (cu |
| gb_pat2:XX084746 | + 2113.00 | 2566.79 | 9.8e-135 | 1826 | AX084746 Sequence 15 from Pat |
| gb_pat1:878957 | + 2077.00 | 2523.53 | 2.5e-132 | 1704 | A78957 Sequence 11 from Patent |
| gb_pat2:XX084732 | + 2075.00 | 2523.46 | 2.5e-132 | 1719 | AX084732 Sequence 11 from Patent |
| gb_pat2:ANPPOX | + 2075.00 | 2521.57 | 2.5e-132 | 1614 | D83139 Arabidopsis thaliana m |
| gb_pat2:XX084742 | + 2062.50 | 2505.16 | 2.7e-131 | 1847 | AX084742 Sequence 11 from Pat |
| gb_pat2:XX084750 | + 2065.50 | 2437.23 | 1.6e-127 | 1785 | AX084750 Sequence 19 from Pat |
| gb_p12:AF160961 | + 2004.50 | 2434.12 | 2.4e-127 | 1924 | AF160961 Cichorium intybus prc |
| gb_p12:AF218052 | + 1889.50 | 2417.42 | 2.0e-126 | 1608 | AF218052 Zea mays protoporphyr |
| gb_pat2:XX084740 | + 1987.00 | 2412.97 | 1.3e-126 | 1811 | AX084740 Sequence 9 from Patent |
| gb_pat2:XX084736 | + 1978.00 | 2402.93 | 1.3e-125 | 1691 | AX084736 Sequence 5 from Patent |
| gb_pat1:878959 | + 1974.00 | 2396.97 | 1.3e-125 | 1923 | A78959 Sequence 13 from Patent |
| gb_p11:AB057749 | + 1965.50 | 2388.97 | 7.9e-125 | 1689 | AB057749 Oryza sativa PPOX-1/F |
| gb_pat2:XX084748 | + 1933.50 | 2347.65 | 1.6e-122 | 1910 | AX084748 Sequence 17 from Pat |
| gb_p14:TX084754 | + 1823.00 | 2214.65 | 4.0e-115 | 1590 | AX084754 Sequence 23 from Pat |
| gb_p12:ATCHR1V4 | + 1726.00 | 2054.12 | 3.5e-106 | 198697 | AF104919 Arabidopsis thaliana |
| gb_p12:655171 | + 1530.50 | 1856.97 | 3.4e-95 | 1838 | E65171 Method for evaluating a |
| gb_p11:AB068635 | + 1524.00 | 1839.54 | 3.2e-94 | 2480 | AB068635 Chlamydomonas reinha |
| gb_p12:AB057771 | + 1524.00 | 1809.90 | 1.4e-92 | 159749 | AB057771 Oryza sativa PPOX-1/F |
| gb_pat2:XX084752 | + 1261.00 | 1532.16 | 4.2e-77 | 1224 | AX084752 Sequence 21 from Pat |
| gb_p12:AF332964 | + 988.50 | 1195.56 | 2.3e-58 | 2060 | AF332964 Polytomella sp. 'Prit |
| gb_p11:AF030179 | + 606.00 | 715.42 | 1.3e-31 | 10265 | AF030179 Chlamydomonas reinha |
| gb_ba1:AF098938 | + 541.00 | 653.59 | 3.6e-28 | 1416 | AF098938 Myxococcus xanthus pr |
| gb_ba3:MMACSGA | + 535.00 | 637.37 | 2.9e-27 | 3909 | L27429 Myxococcus xanthus cs9 |
| gb_pat2:821704 | + 522.50 | 640.26 | 2.0e-27 | 2622 | E21704 Protoporphyrinogen oxid |
| gb_p11:AB025102 | + 522.50 | 629.21 | 8.2e-27 | 1745 | AB025102 Glycine max mRNA for |
| gb_ba1:AF000768 | + 518.00 | 606.19 | 1.6e-25 | 12905 | AF000768 Aquifex aeolicus seq |
| gb_p13:NTY13466 | + 506.50 | 608.79 | 1.1e-25 | 1940 | Y13466 Nicotiana tabacum mrn |
| gb_p11:AB020500 | + 505.00 | 609.13 | 1.1e-25 | 1515 | AB020500 Nicotiana tabacum mrn |
| gb_p11:AF044129 | + 505.00 | 607.23 | 1.4e-25 | 1882 | AF044129 Nicotiana tabacum prc |
| gb_pat2:814138 | + 490.50 | 590.65 | 1.2e-24 | 1662 | E14138 cDNA encoding Arabidops |
| gb_pat2:XX084734 | + 490.50 | 590.26 | 1.2e-24 | 1738 | AX084734 Sequence 3 from Patent |
| gb_ba2:BACHMEHY | + 485.00 | 571.82 | 1.3e-23 | 6624 | M97208 Bacillus subtilis penic |
| gb_ba2:BSY14083 | + 485.00 | 559.83 | 6.0e-23 | 26007 | Y14083 Bacillus subtilis chrd |

gb_ba2:BSUB0006 + 485.00 541.50 6.3e-22 210440 + 1299109 Bacillus subtilis
gb_p12:AF273767 + 467.50 560.82 5.3e-23 2042 + AF273767 Zea mays protoporp
gb_pat2:AX084738 + 467.50 560.74 5.4e-23 2061 + AX084738 Sequence 7 from Pa
gb_p14:STA225108 + 424.00 512.04 2.8e-20 1261 + AJ225108 Solanum tuberosum

seq_name: gb_p11:AF044128

seq_documentation_block:
LOCUS AF044128 1874 bp mRNA PLN 05-JAN-1999

DEFINITION Nicotiana tabacum protoporphyrinogen oxidase PX-1 mRNA, complete

CDs.

ACCESSION AF044128

VERSION AF044128.1 GI:4105185

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

FEATURES

source

CDs

BASE COUNT

ORIGIN

alignment_scores:

Ratio:

Percent Similarity:

Percent Identity:

alignment_block:

US-09-508-418-2 x AF044128

Align seg 1/1 to: AF044128 from: 1 to: 1874

1 MetTrrThrProIleAlaAsnHisProAsnIlePheThrHisGlnse 17

26 ATGACACACACTCCCATGCATCATCATATTTTACCTCACCACTC 75

17 rSerSerSerProLeuAlaPheLeuAsnArgThSerPheIleProPhes 34

76 GCGCAGCATCGCATTCCTTAACCGTACGAGTTTCACTTCCTTCT 125

34 erSerIleSerIysArgAsnSerValAsnGlyTyrPArgThrArg 50

```

126 CTTCAATCTCCAGGCGCAATAGTGCATATTCATATGGCTGGAGAACGCA 175
51 CysSerValAlaIyAspTyrThrValProSerSerAlaValAspGlyG1 67
176 TGTCTCCGTGGCCAAAGATTACACAGTCTCTCCACGCGGTGCGAGCGG 225
67 yProAlaIaIaIeUuaSPcysValIleValG1AlaGlyIleSerGlyL 84
226 ACCGCGCCGAGAGCTGACAGTGTGTATAGTTGGAGCCAGGAATATGTGGCC 275
84 euCysIleAlaG1InValMetSerAlaasnTyrProAsnLeuMetValThr 100
276 TCTGCATTTGCCAGGTGATGTCCGCTAATATACCCCAATTTGATGTACCC 325
101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspG1 117
326 GAGCGAGAGATCGTCCGCGTGGCAACATTAACGACTGTGGAAGAGACGG 375
117 yTyrLeuTyrGluGlyProAsnSerPheGlnProSerAspProMetL 134
376 CTATTTTGGGAGAGAGTCCCAACAGTTTCCAGCCGTCGATCTATGT 425
134 eutHrMetAlaValAspCysGlyLeuUyAspAspLeuValLeuGlyAsp 150
426 TGACTATGCGAGTAGATTGTGATGAAGATGATTTGCTGTGGAGAT 475
151 ProAsnAlaProArgPheValLeuTyrGlyGlyLeuArgProValPr 167
476 CCTAATGGCCCGCTTCTGTTTGTGAAGGGTAAATTAAGCCCGCTCC 525
167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
526 CTCAAACTCAGTATCTCCCTTTTGTGATTTGATGAGCATTCCTTGCA 575
184 ySLeuAlaGlyAlaGlyPheGlyProIleGlyLeuArgProSerProGly 200
576 AGTTGAGAGCGTGTGGTGTGCTCATGTGCTCCCTCACCCTCAGCT 625
201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyGlu 217
626 CATGAGCATAGTTGAGACAGTGTGCTGCTATATCTGTGTGGCGAGAT 675
217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAsp 234
676 CTTGAACGCTGTGATGAGACCATTTTGTGCTGCTTATGCTGTGATC 725
234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrrPlyLLeuGlu 250
726 CCTCAAACTGATATGAAAGCACATTTGGCAAGTTTGGAAAGTTGGAA 775
251 GluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysGlu 267
776 GAAACTGTGTGATCATTTATGGAGAACCTTTTAAACATTAAGGAGAG 825
267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLys 284
826 ATCCAGTAACTGATTAAGCGCCCGCGATCCGCTTACCTAAACCAAAAG 875
284 LysGlnThrValGlySerPheArgGlyGlyLeuArgMetLeuProAspAla 300
876 GACGAGACAGTTGGATCATTCAGGAAGGCTCTCAAAATGCTCCGAGAGCA 925
301 lIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSerSe 317
926 ATCAGTCAAGATTTGGAGAGCAAAATTAACATATCATGAGAACTTTTAG 975
317 rIleThrLysSerGluLysGlyGlyTyrHisLeuThrTyrGluThrProG 334
976 CATTACTTAAGTCAAGAAAAGAGAGATATCACTTACATAGAGACACACAG 1025
334 LuGlyValAlaSerLeuGlnSerArgSerIleValMetThrValProSer 350

```

```

1026 AAGAGTAGTTCTCTTCAAAAGTCGAAGCATTTGTATGATGCTGGCATCC 1075
351 TyrValAlaSerAsnIleLeuArgProIleUaSerValAlaAlaIaAspAl 367
1076 TATGTAGCAAGCAACATATTAGCTCTCTTGTGGTGGCGAGAGATGC 1125
367 aluSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
1126 ACTTTAAATTTCTACTATCCCCAGTGTGAGCAGTGCACAAATTTCTATTC 1175
384 roGlnGluAlaIaIaArgAspGluArgLeuValAspGlyLLeuUySgly 400
1176 CTCAGAACCTATTCGTATGAGCGCTCGTTGATGCTGAACATAAGGGA 1225
401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrI 417
1226 TTTGGCAGATTGCATCCAGTACACAGGAGTGGAAACCTAGGAACGAT 1275
417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
1276 ATATAGTTCATCTCTCTTCCCTAACCGTCCCAAAAGGTGCGGTCTAC 1325
434 euLeuAsnTyrIleGlyAlaIaLysAsnProGluIleUaSerLysThr 450
1326 TCTTGAACATACTGGAGAGCAAAAATCTGAATTTGTCTAAGACG 1375
451 GluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMetLeuI 467
1376 GAGAGCAACCTGTGGAAATGATGTATCGTACCTCAGAAAATCTTAT 1425
467 eLysProLysAlaGlnAspProLeuValAlaGlyValArgValTrrProG 484
1426 AAACCCAAAGCTCAAGATCCTTGTGTGGTGTGCGAGATAGGCCAC 1475
484 lAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
1476 AAGCTATCCACAGATTTTGTGTGTCATCTGATACGTATGACTGCA 1525
501 LysAlaAlaMetAsnAspAsnGlyLeuGlyLeuPheLeuGlyGlyAs 517
1526 AAGCTGCTATGATATATATGAGCTTGAGAGGCTGTTCTTGGGGGCTNA 1575
517 nTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGlu 534
1576 TTATGTGTCAGGATGATGATGAGGAGTGTGTTGAAGTCTGTATGAAG 1625
534 alAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
1626 TTGCATCCGAGGTAAACAGATTTCTGTCTGCTATGATGCATAA 1669
seq_name: gb_p11:AB020501
seq_documentation_block:
LOCUS AB020501 1647 bp mRNA PLN 06-FEB-1999
DEFINITION Nicotiana tabacum mRNA for plastidal protoporphyrinogen oxidase,
complete cds.
ACCESSION AB020501
VERSION AB020501.1 GI:3929921
KEYWORDS plastidal protoporphyrinogen oxidase.
SOURCE Nicotiana tabacum (strain:Samsun NN) cDNA to mRNA.
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE
1 (bases 1 to 1647)
AUTHORS Che,F. and Watanabe,N.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1998) to the DDBJ/EMBL/GenBank databases.
Fang-Sik Che, Nara Institute of Science and Technology, Graduate
School of Biological Sciences, Takayamachou 8916-5, Ikoma, Nara
630-0101, Japan (E-mail: fische@bs.aist-nara.ac.jp,
Tel:81-743-72-5453, Fax:81-743-72-5459)
REFERENCE
2 (sites)

```

AUTHORS Watanabe, N., Che, F. S., Iwano, M., Takayama, S., Nakano, T., Yoshida, S. and Isogai, A.
 TITLE Molecular characterization of photomixotrophic tobacco cells resistant to protoporphyrinogen oxidase-inhibiting herbicides
 JOURNAL Plant Physiol. 118 (3), 751-758 (1998)
 MEDLINE 99026276
 FEATURES
 source Location/Qualifiers
 1..1647
 /organism="Nicotiana tabacum"
 /strain="Samsun NN"
 /db_xref="taxon:4097"
 /note="Photomixotrophic cultured cells"
 1..1647
 /gene="NtPROX1"
 1..1647
 /gene="NtPROX1"
 /codon_start=1
 /product="plastidial protoporphyrinogen oxidase"
 /protein_id="BA34713.1"
 /db_xref="GI:392922"
 /translation="MTTPPIAHNPNIPTHOSSSPLEFLNRTSEIPSSISKRSVNC
 NGMRTCSVANDYTPSSAVDGPAAELDCYIVAGISGLCIAVMSANPNLVTEA
 RDRAGNITVEEDYIWEEGNSFOPSPMLTVAIDGLDIDLVLGPPNAPRYLWK
 GKLRPVSKLIDIAFPDLMSTIPGKLRAFGAIGIRPSPGHESEVEDPVRNIGEVF
 ERLIEPSSGVAGDPSKSKMAKAPGKWKLEETGSLTGTFKAIKRSSTPRAPD
 PRLRPKQITVGSERFKGLRMLPDALISARLSKLSLWSLSTITSEKGYHLTEPE
 GVSLQSRSIWTVPSYVANSILRPLSYAADALSNFYPPGAVTISYPOEATDER
 LVDELKGFQGLHPRTQGVETIGTYSSSLPPNAPKGRVLLNVIIGAKNPETLSKT
 ESOLVEVDRDLRKMLIKPKADPLVGVVPMPOALIPFLVGHLDLTSTAKAANDNG
 LEGFLGNGNYSGVALGRCVGEVASEVTEGFLSRAYK"
 BASE COUNT 437 a 349 c 415 g 446 t
 ORIGIN
 alignment_scores:
 Quality: 2824.00 Length: 548
 Ratio: 5.172 Gaps: 0
 Percent Similarity: 99.635 Percent Identity: 99.453
 alignment_block:
 US-09-508-418-2 x AB020501 ..
 Align seg 1/1 to: AB020501 from: 1 to: 1647
 1 MetThrThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnse 17
 1 ATGACAAACAACCTCCATCGCCAAATCCAAATATTTCACATCCACGATC 50
 17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheS 34
 51 GTGCTATCGCCCATTTGGCATTTCTTAAACCGTAGAGTTTCATCCCTTTCT 100
 34 exSerIleSerLysArgAsnSerValAsnGlySerGlyTyrParGThrArg 50
 101 CTCATATCTCCAGGCCAATAGTGCATTTGCATTTGGCTGGAGAACACCA 150
 51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAspGlyG 67
 151 TGTCTCGTTGGCCAAAGATTACACAGTTCTCTCGTCAGCGGTCAGCGCGG 200
 67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
 201 ACCCGCGCGGAGCTGGACTGTGTATTAGTGGAGCGAGGAATTAGAGGCC 250
 84 euGlyIleAlaGluValMetSerAlaAsnTyrProAsnLeuMetValTyr 100
 251 TCTGCATTCGCGAGGTGATGTCGCGTAATTACCCCAATTTGATGTGAAC 300
 101 GluAlaArgAspArgAlaGlyLysAsnIleThrThrValGluArgAspG 117
 301 GAGCGAGAGATCGTCCGCGGCAACATACAGCTGTGGAAGAGAGACGG 350
 117 yTyrLeuTyrPgluGluGlyProAsnSerPheGlnProSerAspPrometL 134

351 CATTTGTGGAGAAAGATGCCAACAGTTTCCAGCCGTCGATCTTATGT 400
 134 euThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyasp 150
 401 TGCATATGCGACATAGATTTGTGATGAAGATGATTTGGTGTGGAAAT 450
 151 ProAsnAlaProArgPheValLeuTyrPlyGlyLysLeuArgProValPr 167
 451 CCTAATGGCCCGCTTTCGTTTGTGGAAAGGTAATTAAGGCCGCTGCC 500
 167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyL 184
 501 CTCAAAACCTCACTGATCTTGCTTTTGTGATTTGATGAGCATTCCTGCA 550
 184 yLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
 551 AGTTGAGAGCTGTTTTGTTGCTGATGGCTCCGCTTCACCTCCAGGT 600
 201 HisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGlyLys 217
 601 CATGAGGATCATGTTGAGAGAGTTCGTCGCTATCTTGGTGGCGAAGT 650
 217 lPheGluArgLeuIleGluProPheCysSerGlyValIleValGlyasp 234
 651 CTTTGACCTTGATAGAACCATTTTGTCTGTTATGCTGTGTGATC 700
 234 rSerLysLeuSerMetLysAlaAlaPheGlyLysValTyrPlySerLeu 250
 701 CCTCAAACTGAGTATGAAGAGCATTTGGGAAAGTTTGAAGTTGGA 750
 251 GluThrGlyGlySerIleIleGlyLysThrPheLysAlaIleLysGluar 267
 751 GAACCTGGTGATGATTTATGAGAGAACCTTTAAAGCAATTAAGAGAG 800
 267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProlysG 284
 801 ATCCAGTACACCTTAAGCGCCCGGATCCGGTTTACCTAAACCAAG 850
 284 lGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
 851 GACACACAGTTGATCATTCACAGAAAGGCTTCAGATCTCCGGATGCA 900
 301 lLeSerAlaArgLeuGlySerLysLeuLysLeuSerTyrPlyLysSerSe 317
 901 ATCAGTGCAGATTTGGAAAGCAATTAACATCATCATGGAAGCTTTAG 950
 317 rIleThrLysSerGlyLysGlyLysThrHisLeuThrTyrGluThrProG 334
 951 CATTTACTAAGTCAGAAAAGAGGATATCACTTGACATACAGACACAG 1000
 334 lGluValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
 1001 AAGGAGTGTTCCTTTCAAAAGTCGACACATTCATGACGTGCTCATCC 1050
 351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
 1051 TATGTAGCAAGACATATTTAGTCCCTCTTCGTTGCGGACAGAGATGC 1100
 367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
 1101 ACTTTCAAAATTTTACTATCTCCCGCAGTTGGAGCAGTACACATTTCAATC 1150
 384 rGlnGlnAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
 1151 CTCAGAGAGCTATTCGTATGAGCGTCTGTTGATGTGTAACCTAAAGGA 1200
 401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrI 417
 1201 TTTGGGCGATTGCATCCACGTACACAGGAGTGAAGACACTAGAGACGAT 1250
 417 eTyrSerSerLeuPheProAsnArgAlaProLysGlyValArgAlaLeu 434

```

1251 ATATAGTTCATCACTCTTCCCTAACCGTGCCCAAAAGCTGGGGTGTAC 1300
434 euleuasnTYRllieglyAlaLysasnProLulleuSerLysThr 450
1301 TCTTGACACTGATTGGAGGACAAAATCTGAATTTGTCTAAGACG 1350
451 GluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMetLeuI 467
1351 GAGGCGCAACTTGTGGAGAGTAGTTGATCGTCACTCGAAAAATGCTTAT 1400
467 eLysProLysAlaGlnAspProLeuValGluValAlaArgValTyrProG 484
1401 AAAACCCAAAGCTCAAGATCTCTTGTGGGAGTGTCGAGTATGGCCAC 1450
484 lnaAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
1451 AAGCTATCCACACAGCTTTTGGTGGTATCGATCGATAGCTAAGTACGCA 1500
501 LysAlaAlaMetAsnAspAsnGlyLeuGlyGlyLeuPheLeuGlyLys 517
1501 AAAGCTGCTATGAATGATATATGGCTTGAAGGCGTCTTCTGGGGGTAA 1550
517 nTYrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluV 534
1551 TTATGTGTCAAGTGTACATGGGAGGTGTGTGAAGGTGCTTATGAAG 1600
534 aAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
1601 TTGCATCCGAGGTAAACAGATTCTGTCTGGTATGATCAAA 1644
seq_name: gb_p13:NTY13465

```

```

seq_documentation_block:
LOCUS NTY13465 1892 bp mRNA PLN 08-SEP-1997
DEFINITION Nicotiana tabacum mRNA for protoporphyrin IX oxidase, ppxI gene.
ACCESSION Y13465
VERSION Y13465.1 GI:2370332
KEYWORDS oxidase; ppxI gene; protoporphyrin IX; tetrapyrrole synthesis
enzyme.
SOURCE common tobacco.
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Asteridae; eusterids
I; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1892)
AUTHORS Lermontova,I., Kruse,E., Mock,H.P. and Grimm,B.
TITLE Cloning and characterization of a plastidial and a mitochondrial
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 94 (16), 8895-8900 (1997)
MEDLINE 97385200
REFERENCE 2 (bases 1 to 1892)
AUTHORS Inna,L.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-1997) L. Inna, Institute of Plant Genetics (IPK),
Dept. of Molecular Cell Biology, Corrensstr.3, 06466 Gatersleben,
FRG
REMARK revised by author
FEATURES
Source location/Qualifiers
1..1892
/organism="Nicotiana tabacum"
/strain="SRI"
/db_xref="taxon:4097"
/clone_lib="Lambda ZAP II tobacco library"
29..1675
/gene="ppxI"
29..1675
/gene="ppxI"
29..1675
/gene="ppxI"
/EC_number="1.3.3.4"
/function="tetrapyrrole biosynthesis"
/codon_start=1
/product="protoporphyrinogen oxidase"
/protein_id="CAA73865.1"
/db_xref="GI:2370333"

```

```

/db_xref="SWISS-PROT:024163"
/translation="MTTTPRIANPNIFTHQSSSSPLAFINRTSIFPSSIRKSNVNC
NGMRPCSAVKDQVTPSSAVDDGPAEELDCVIGISGLIAQVMSANYPLMAYTEA
RDRAGSNITTVENDQGLMEGPNRPPSPMPTMVDGLKDLIVIGDNPAPRYLAK
GKLRPVPSKITDLAFDLMSTPCKLAKAGCAIGLRSPFGHESVDEQPYRRLLGGEVF
ERLIEPFGSVVAGDPSKLSMKAFGKWKLEDTGSLIGTFKAIKESSTPKAPRD
PRLPKRGQVVGSEFRKGLNMLPDALISARLGSKLSKLSLSTIKSEKGYHLTYTPE
GVSLOSRSIVMTVPSPYSANILRPLSAADALSNFYYPVGAVTYTIPQDAIRDER
LYDQELKFEQSLHPTROGVEETLGTIVSSLSFPRARAKGVLLNIGCAKNDPEILSKT
ESQLVYVPDRLERKMLIKPKADPLVGVYVMPQALPQFLVGHLDTLSTAKAAMDNG
LEGFELGAGNVSGVALGRCEVGEAYEVASEVTGLSYAK"
BASE COUNT 516 a 383 c 467 g 526 t
ORIGIN

```

```

alignment_scores:
Quality: 2821.00 Length: 548
Ratio: 5.167 Gaps: 0
Percent Similarity: 99.635 Percent Identity: 99.270

```

```

alignment_block:
US-09-508-418-2 x NTY13465

```

```

Align seg 1/1 to: NTY13465 from: 1 to: 1892

```

```

1 MetThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSe 17
|||||
29 ATGACACAGACATCCCATGCGCAATCATCTATATTTACACTACACAGATC 78
|||||
17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPhe 34
79 GTCGTCATCGCCATGTGGCATTCCTTAACCGTACGAGTTTCATCCCTTCT 128
|||||
34 eSerIleSerLysArgAsnSerValAsnCysAsnGlyTyrPargThrArg 50
129 CTTCATATCTCCAGCGCAATATGTCATATGCAATGGCTGGAGAACAGCA 178
|||||
51 CysSerValAlaLysAspTyrThrValProSerSerAlaValAlaSpGly 67
179 TGCTCCGTTGCCAAGATTACACAGATTCCTTCACGCGGTGACGCGCG 228
|||||
67 yProAlaAlaGluLeuAspCysValIleValGlyAlaGlyIleSerGly 84
229 ACCCGCGCGGAGCTGAGCTGTGTTATAGTTGAGCAGGAAATTAAGTGC 278
|||||
84 eucysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
279 TCTGCTATGCGCAGGTGATGTCGCTAATTACCCCAATTTGATGGTAACC 328
|||||
101 GluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGluArgAspG 117
329 GAGCGCAGAGATCGTCCGCTGCAACATACGACTGTGGAAGAGACGCG 378
|||||
117 yTyrLeuThrPgluGluGlyProAsnSerPheGlnProSerAspProMet 134
379 CTAATTTGTGGAGAGAGGTCCTCAACAGTTTCCAGCGCTCCGATCTATGT 428
|||||
134 euthrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlyAsp 150
429 TGACTATGCGCAGTAGATTGTGATTAAGAGATTTGGTGTGGAGAGT 478
|||||
151 ProAsnAlaProArgPheValLeuTyrPlyGlyLysLeuArgProValPr 167
479 CCAATATGCGCCCGCTTGTGTTGTGGAAGGTAAATTAAGCCCGCTGCC 528
|||||
167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerIleProGly 184
529 CTCMAAACCTCACAGATCTTGTCTTTTATTTGATGAGACATCTCTGGCA 578
|||||
184 yLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
579 AGTTGAGAGCTGTTTGGTGGCATTGGCTCGCCCTTCACTCCAGGT 628
|||||

```


201 HisGluGluSerValGluGluInpHeValAArgArgAsnLeuGlyGlyGluVal 217
 217 1PheGluAArgLeuIleGluProPheCysSerGlyValIlyrValGlyAsp 234
 679 CTTTGAACGGTTGATGAACCATTTTGTGCTTATGCTGGGATC 728
 234 roserLysSerMetLysAlaAlaPheGlyLysValTrpLysLeuGlu 250
 729 CTTCAAACTAGATGAAAGCAGCATTTGGGAAAGTTGGAAAGTTGGAA 778
 251 GluThrGlySerIleIleGlyGlyThrPheLysAlaIleLysGluAr 267
 779 GAACAGGTGGTGATGATTTATGAGAGACCTTTAAAGCAATTAAGAGAG 828
 267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
 829 ATCCAGTACACCTAAAGCGCCCGCATCCGCTTACCTAAACCAAAAG 878
 284 LysGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
 879 GACAGACAGTTGATCATTCAGAGAGGCTCTCAGAAATGCTGCCGATGCA 928
 301 IleSerAlaArgLeuGlySerLysLysLeuLysSerTrpLysLeuSerSe 317
 929 ATCAGTGCAGATTTGGGAAGCAATTAACATCATGAGACGTTCTAG 978
 317 rIleThrLysSerGlyLysGlyGlyTyThrIleThrIlyrGluThrProG 334
 979 CATTACTAACTCAGAAAAGAGGATATCACTTGACATCAGAACACACAG 1028
 334 LysGlyValIleSerLeuGlnSerArgSerIleValMetThrValProSer 350
 1029 AAGCAGTAGTTCTCTTCAAGTCGAAAGCTTCTCATGACTGCGCATCC 1078
 351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
 1079 TATGTACAGCAACANATATAGCTCTTGGTGGCCGACAGATGC 1128
 367 AleuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
 1129 ACTTCAAAATTTCTACTATCCCCAGTTGGAGCAGCACATTAATCATATC 1178
 384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGlyLeuLysGly 400
 1179 CTCAAGAGGCTATTGATGAGGCTGCTGTATAGGTGAACTAAGAGGA 1228
 401 PheGlyLysLeuLysProArgThrGlnGlyValGluThrLeuGlyThrI 417
 1229 TTGGGCAAGTTGCATCCAGCTACACAGGAGTGGAACACTAGAACGAT 1278
 417 eTyrrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuL 434
 1279 ATATAGTTCACTACTCTTCCCTAACCGTGCCCAAAAGGCGGCTGTAC 1328
 434 euLeuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
 1329 TCTTGAACCTATTGGAGAGCAAAAATCTCTGAATTTTGTCTTAAGACG 1378
 451 GluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMetLeuI 467
 1379 GAGAGCAACCTTGTGGAAGTAGTTGATCGTACCTCGAATAATGCTTAT 1428
 467 eLysProLysAlaGlnAspProLeuValAlaGlyValArgValThrProG 484
 1429 AAAACCCAAAGCTCAAGATCTCTTGTGTGGGTGCGAGTATGGCCAC 1478
 484 InAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
 1479 AAGCTATCCCAAGATTTTGTGTTGCTATCATCGATACGTAAGTACTGCA 1528
 501 LysAlaIleAlaMetAsnAspAsnGlyLeuGlyLysLeuPheLeuGlyLys 517

1529 AAAGTGTATGAATGATTAATGAGGCTTGAAGGCTGTTCTTGGGGTAA 1578
 517 nTyrrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluVal 534
 1579 TTATGTGACAGGTGATGAGCTTGGGAGGCTGTCTTGAAGGCTCTTATGAG 1628
 534 AlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
 1629 TTGCATCCGAGGTACAGCATTTCTGCTCGTATGCATACAA 1672
 seq_name: gb-pl4:STA225107
 seq_documentation_block:
 LOCUS STA225107 1976 bp mRNA 09-FEB-1999
 DEFINITION Solanum tuberosum (cultivar Bintje) chloroplastic
 protoporphyrinogen IX oxidase.
 ACCESSION AJ225107
 VERSION AJ225107.1 GI:3093409
 KEYWORDS protoporphyrinogen oxidase.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
 I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 1976)
 Johnston, D.J.
 TITLE Direct Submission
 JOURNAL Submitted (23-APR-1998) Johnston D.J., Biologie moleculaire, RAC,
 Changins, 1260, Nyon, SWITZERLAND
 2 (bases 1 to 1976)
 Johnston, D.J., Droz, E., Rochaix, J.D. and Malno, P.
 Cloning and Characterization of Potato cDNAs Involved in
 Tetrapyrrole Biosynthesis: ferrochelatase (Accession No. AJ005802),
 chloroplastic protoporphyrinogen IX oxidase (Accession No.
 AJ225107), and mitochondrial protoporphyrinogen IX oxidase
 (Accession No. AJ225108)
 JOURNAL Plant Physiol. 118, 330-330 (1998)
 FEATURES
 location/Qualifiers
 1..1976
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 1..1674
 /EC_number="1.3.3.4"
 /function="tetrapyrrole biosynthesis"
 /product="protoporphyrinogen oxidase"
 /protein_id="CAA12400.1"
 /db_xref="GI:3093410"
 /db_xref="SPRREMBL:O64384"
 /translation="MTTAAVANHPSIFTHRSPLPSPSSSSSPFLFLNRTNFIPEYS
 TSKRNSVNCNGMRTRCSVAKDYVPPSEVDNGNFPFLDCVYGAGISGLCIKAVISAN
 YPLNMTYEADRRAGNITTYERDGYLMEBSPNFQSPDMILNANDCGLKADLYISDP
 DAREFVLMKOKLPEVPGKLTDPEDFDMSTPGKLRAGFALIGRSPGPIESVDEQFV
 RNRNAGVFERLLEPFCGSVYAGDPSKLIKRAFGKVMKLEDTGGSIIIGETKAIKER
 SNMKPRDRRLPLPGQGVYSEFKGLRMLPDALICRLSLKYSLSWKLSSIRKSEKGG
 YLLFYEPEGVSLRSRYVMYVSVYINILRLSVAAADALSFFYPPVAAVTISY
 PORAIRDERLYDEGLKQHLHRSQGVETIGITSSSLFPNRAIPKGRVLLNLYIGGA
 TNEIYKTSQSQLEAVDRDGLKMLKPKRAQDPFYVGVYWPQALIPQFLYGHLDLTGT
 AKTALSDNGIDGLFELGNGVSVGLRCVBAEYIASVETGFLSQAYR"
 1675..1976
 3' UTR
 BASE COUNT 518 a 394 c 485 g 579 t
 ORIGIN
 alignment_scores:
 quality: 2565.50 Length: 557
 Ratio: 4.850 Gaps: 2
 Percent Similarity: 94.973 Percent Identity: 88.869
 alignment_block:
 US-09-508-418-2 x STA225107

Align seg 1/1 to: STA225107 from: 1 to: 1976

1 MethrthrhPrroIleAlaSnHsPProsnIlePmthRHsGln... 16
1 ATGACAAACAAAGCGCGCTCGCACCACTCTTACATTTTCACTCAACGGGTC 50
17SerserSerserPro...LeuAlaPheL 25
51 GCCCGCTCCGCTCCGCTCGCTCCCTCCATCCGCGCTCATTTT 100
25 euASnArgrHsrserPheIleProPheSerSeriIleSerLyArASnSer 41
101 TTAACCGTACGAATTTTCATTCTCTTACTTTTCCACCTCCAAACGGCAATAGT 150
42 ValASnCyASnGlyTrpArgThrArgCySerValAlaLysAspYrTrh 58
151 GTCATTTGCATGCGCTGGAGAACAGATGTCCTCGTCCGAAGATTTATC 200
58 rValProSerSerAlaValAlaSpGlyGlyProAlaAlaGluLeuAspYsY 75
201 AGTTCCCTCCCTCGGAAGTGCAGCGTAATAGTCCCGAGGTGGATTGTG 250
75 alIleValGlyAlaGlyIleSerGlyLeuCyIleAlaGluValMetSer 91
251 TCGAATGTGGAGCAGGAATTAGTGACCTCGATTGCTAAAGTCATTTCG 300
92 AlaAspTrpProAsnLeuMetValThrGluAlaArgAspArgAlaGly 108
301 GCTAATATCCCAATTGTAGTGGAGCGAGCGAGGATCTCGCGGTGG 350
108 yASnIleThrThValGlaIaArgAspGlyTYrLeuTrpGluGlyPro 125
351 AAACATTAACGACGCTGGAAAGAGATGATCTTAATGGGAAGAAGTCTCA 400
125 snSerPheGlnProSerAspPrometLeuThrMetAlaValAspCyGly 141
401 ACAGTTTCCAGCCTTGCGATGCTATGTTGACAAATGGCTGTAAGTTGGTA 450
142 LeuLyASpAspLeuValLeuGlyLysAspProAsnAlaProArgPheValLe 158
451 TTGAAGATGATTGGTGTGGGAAATCCTGATCGCTCGCTTGTGCTT 500
158 uTrpLyGlyLysLeuArgProValProSerLysLeuThrAspLeuPro 175
501 GTGGAAAGATAACTAAGGCGCTGTTCCGCGCAACCTCACTGATCTTCCG 550
175 hepASpAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyPro 191
551 TCTTTGATTGATGAGATATCCTCGCAACGCTCAGAGCTGTTTGGTGGC 600
192 IlleGlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPh 208
601 ATTTGGCTTCGGCTTACCTCCAGGTTATGAGAACTAGTTAGCAGT 650
208 eValArgArgAsnLeuGlyGlyValAlaPheGluArgLeuIleGluPro 225
651 CGTGGCTGTAATCTTGGTGGAGAGTCTTTGAACGTTTATTTGAACCAT 700
225 heCySerSerGlyValYrValGlyAspProSerLysLeuSerMetLysAla 241
701 TTTGTTCTGGTGTTCACGCGGTGACCCCTCAAAATGATTATGAAGA 750
242 AlaPheGlyLysValTrpLysLeuGluGluThrGlyLysIleIleLeu 258
751 GCATTTTGGAAAGCTGGAGAGCTAGAACAAACCTGTGTGATTTTGG 800
258 yGlyThrPheLysAlaIleLysGluArgSerSerThrProLysAlaPro 275
801 GGAACCTTTAAAGCAATTAAGAGAGATCCAGTAACCTTAACCGCCTC 850
275 rGAspProArgLeuProLysProLysGlyGlnThrValGlySerPheArg 291

| | | |
|--|---|------|
| 851 | GTATATCGCGGTTTACCACACACCAAAAGGACAAACTGTTGGATCATTTAGG | 900 |
| 292 | lysGlyIleuAtrMetLeuProAspAlaIleSerAlaArgLeuGlySerIy | 308 |
| 901 | AAAGGCTGTGAATGCTCCCGAATGCACAAATTTGCGAAAGACTGGGAACAA | 950 |
| 308 | sIeuIyIleuSerTrpIyIleuSerSerIleThrIySergIuIySgIyG | 325 |
| 951 | AGTAAACATCATCATGGAAGCTTCTAGCATTTACAACTCGAAAAAGAG | 1000 |
| 325 | IyTrpHisIleuThrTyrgIuThrProGluGlyValIalSerLeuGlnSer | 341 |
| 1001 | GAATCTCTTGACATACAGACACACAGAGAGTAAgTTCTCTGCGCAAGT | 1050 |
| 342 | ArgSerIleValMetThrValProSerTyTrValAlaSerAsnIleLeuAr | 358 |
| 1051 | CGAAGATTGTCATGACTGCTTCCATCTCTATGTAGCAACGACATATTACG | 1100 |
| 358 | gProIeuSerValAlaAlaAlaAspAlaLeuSerAsnPhetyTyTrPro | 375 |
| 1101 | CCCTCTTTTCGGTGCCTGCAGACAGATGCACCTTCAAGTTCTACTATCCCC | 1150 |
| 375 | rovalGlyValAlaIthrIleSerTyProGlnGluAlaIleArgAspGlu | 391 |
| 1151 | CGATACACACAGTACACATTTCAATCTCTCAAGAGGCTATTCTGTATGAG | 1200 |
| 392 | ArgLeuValAspGlyGluLeuIleuGlyIlePheGlyGlnIleuHisProArgTh | 408 |
| 1201 | CGCTGTGGTTGATGGAGCTAAAGGATTTGGGAGTGGCATCCACGTTTC | 1250 |
| 408 | rgIuGlyValAlaGuthrIleuGlyThrIleTySerSerSerIeuPheProA | 425 |
| 1251 | ACAGGAGAGGGAACACATCAGGAAACAAATATATATCTATCATCTTTCCCA | 1300 |
| 425 | sArGAlAProIySgIyArgValIeuIeuIeuAsnTyrrIleGlyGlyAla | 441 |
| 1301 | ACCGTGCTCCAAATGCGCGGGTGTACTCTTGAATCATCATGGAGAGCA | 1350 |
| 442 | IySAsnProGluIleIeuSerIySthrGluSergIuIeuValGluValVa | 458 |
| 1351 | ACAAATACTGAAATTTGTGTCAAGACGAGACGACCACTGTGGAGCACT | 1400 |
| 458 | IAspArgAspIeuArgIySmetIeuIleIySProIyAlaGlnAspProL | 475 |
| 1401 | TCACCGTACCTCGAATAAATGCTTATAAAAACCAAGCAACAGATCCCT | 1450 |
| 475 | euValIalGlyValArgValItrProGlnAlaIleProGlnPheIeuVal | 491 |
| 1451 | TTGTTACGGGTGTGGAGTATGGCCACAACGTAATCCACAGTTTTTGGTC | 1500 |
| 492 | GlyHisIleuAspTrpIleuSerThrAlaIyAlaAlaIleMetAsnAspAsnG | 508 |
| 1501 | GGACATCTGGATACACTAGTACGCAAAAAACGCTCTAAGGATAAATGG | 1550 |
| 508 | yIeuIuIyIeuPheIeuGlyGlyAsnTyrrValIleSrgIyValAlaIleuG | 525 |
| 1551 | GGTTACGGGCTATTCTCTGGGGTATATTATGTCTGGGTGATGACATTGG | 1600 |
| 525 | IyArgCysValGluGlyAlaTyrgIuValAlaIleSergIuValIthrGlyPhe | 541 |
| 1601 | GAAGGTGTGTGAAGGCTTATGAATAATACATCATGAGTAACCTGATTT | 1650 |
| 542 | IeuSerArgTyrrAlaTyrrIyS | 548 |
| 1651 | CTGTCTCAGTATGATACAAA | 1671 |
| seq_name: gb_pat2:AX084746 | | |
| seq_documentation block: 1826 bp DNA PAT 09-MAR-2001 | | |
| LOCUS AX084746 | | |
| DEFINITION Sequence 15 from Patent WO0112825. | | |
| ACCESSION AX084746 | | |
| VERSION AX084746.1 GI:13274999 | | |

KEYWORDS upland cotton.
SOURCE Gossypium hirsutum
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; kosisidae; eurosids II; Malvales; Malvaceae; Gossypium.
REFERENCE 1 (bases 1 to 1826)
AUTHORS Johnson, M. A., Volrath, S. L., Helfetz, P. B. and Law, M. D.
TITLE Herbicide-tolerant protoporphyrinogen oxidase
JOURNAL Patent: WO 0112825-A 15 22-FEB-2001;
Syngenta Participations AG (CH)
FEATURES
source Location/Qualifiers
1..1826
/organism="Gossypium hirsutum"
/db_xref="taxon:3635"
31..1650
/note="cotton protox-1 coding sequence"
/codon_start=1
/protein_id="CAC34014.1"
/db_xref="GI:13275000"
CDS
translation="MTALIDLSLRSSPSVPSIPHHPHPRFRKPKLRLCSLAEGP
TSSSKIDGSGSSIAQCYVGGTSGCTAGALATKHHDVAENVITVEARPGVNT
TVERDGYLMESGPNPSFOPSDPILTMADVSGLKLDIVLADPNAPRVIMEGLRVPK
PTDLPFDLMSIAGKLRAGFGAIGIRPPGVESEVEFVRNLGAVEFERIIPFCS
GVGADPEKLSMKAFGRVWKLEIGTSIIIGTERTIOERNKTEKPPEDPLRPKQ
TVGSEPKGLTMLPEALIANSLGNSVLSWKLSTIKLGGNKEIETFECEGVLSQSS
VYMTIPSHVANSILHPLSAAADALSOFYPPVAVSYVSKAIRMKCLIDGELKGF
GOLHRSQGITIGITVSSSLPPNAPSGRVLLNVTGCAINTGILSTBEELVEAVD
RDLKMLNPNAKDPDIVGVYRWPKAIRPOLYGHLDLDSAKMALRDSGFHGLFLGN
IVSGVALRCVGAIEVAEVEKEFISQYAYK"
BASE COUNT 471 a 399 c 439 g 517 t
ORIGIN

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x AX084746 ..

Align seg 1/1 to: AX084746 from: 1 to: 1826

```

47  TTParGThrArGcSerValAlaLysAspTyrThrValProSerSerAl 63
   ::::: |||||||:::|||||:::|||||:::|||||:::|||||
130 TTCAAGCTCGGATGCTCCCTCGCGAGGCTCCACGATTCCTCATCTAA 179
   :::|||||:::|||||:::|||||:::|||||:::|||||
63  aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
   :::|||||:::|||||:::|||||:::|||||:::|||||
180 AATCGAGCGGGGAGAAATCATCATCGCGGATTCGTCATCGTTGAGGTG 229
   :::|||||:::|||||:::|||||:::|||||:::|||||
80  lytleSerGlyLeuCysTlleAlaGlnIleAlaMetSerAlaAsnTyr.... 94
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
230 GTATCAGTGGAGCTTTCATGCTCAAGCTCTCGCCACCAAGCACCGTGAC 279
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
95  .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
280 GTCCGTTCCATATGATGTGACGAGGAGCCAGACCGTGTGTGGTGCAC 329
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
109 ntleThrThrValGluArgAspGlyTyrLeuTTPGluGluGlyProAsn 126
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
330 CATCACAACCGTGTGAGAGAGATGATATCTGTGGGAAGAAGGCCCAACA 379
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
380 GTTTTCAGCCCTCCGATCTATTCTTAACCATGGCCGTGGATGTGATG 429
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
143 LysAspAspLeuValIleuGlyAspProAsnAlaProArgPheValLeu 159
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
430 AAGGACGATTTGTTAGTGACCTTAATGACACCGGATTTGTAGCATG 479
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176

```

```

480 GGAGCGAAACATTAAGCCGTGTGCTCCACAGCCAAACCATTCGCGTTT 529
   ||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
176 hAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
530 TTGATTTGATGAGCATTCGTGGAAGAACTTAGGGCTGGGGTGGGCTAT 579
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
193 GlyLeuArgProSerProProGlyHisGluGlnIleValIleGluInPhe 209
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
580 GGCATTCGGCCCTCCCGCTCCGGGTATATGAAGAAATCGGAGAGATTGT 629
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPhe 226
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
630 GCGCGGTATCTTGTGCTGAGAGTTTGTGAACGCTTATTAACCATATT 679
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
226 ySerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIle 242
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
680 GTTCAGGTGTATTATCAGAGGATCCTTCAAAATTAAAGCATGAAGCACA 729
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
243 PheGlyLysValTyrPheGluGluGluTyrGlySerIleIleGlyGly 259
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
730 TTGGAGAAGATGAGAGCTAAGAGATGTGGCAGCATCATCTGTGTG 779
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
780 CACTTCAAGACAAATCCAGGAGAGAAATAGACACCTAAGCCACCAAG 829
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
276 sPProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
830 ACCCGGTGTGCCAAACCGAAGGGCCCAACAGTTTGATCTTTAGGAAG 879
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
880 GGACTTACCATGTGCTGAGGACCAATGTCTACAGTTGGTAGCAATGT 929
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
309 uLysLeuSerTyrPheLysLeuSerSerIleThrLysSerGluGlyGly 326
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
930 AAATATATCTTGAAGCTTTCAGATTAACCAATTTGGGCAATGGAGGT 979
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
326 yHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
980 ATACTTGAATTTGAACACCTGAAAGAAATGTATCTTCAGAGTAGA 1029
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArg 359
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1030 AGTGTGTATGACCATTCATCCATGTTGCCAGTAATCTTTCATCC 1079
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
359 oLeuSerValAlaAlaAspAlaLeuSerAsnPheTyrTyrProProy 376
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1080 TCTCTCGGCTGTGCTGCGAGATGCAATATCCCAATTTATTAATCTCCAG 1129
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
376 aLysLeuValThrIleSerTyrProGluGluAlaIleArgAspGluArg 392
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1130 TTGCATCAGTCAAGCTCTCTATCCAAAAGAACCTTGGAAAAGAAATGT 1179
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
393 LeuValAspGlyGlyLeuLysGlyPheGlyGlnLeuHisProArgThrG 409
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1180 TTGATGTAGTGTGAAGCTTAAGGGGTTTGGCAGTTGCACCGCAGCA 1229
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsn 426
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1230 AGGAATTGAAACTTAAAGGACATATACAGTTCAATCATCTTCCCAATC 1279
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
426 rGluProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1280 GAGCTCATCTGGCAGAGGTGTGCTTGAACATACTAGAGAGAGCTACC 1329
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
443 AsnProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAs 459
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
1330 AACACTGGAATTTTTCCAAAGACTGAAGGGGAACTTGTGAGAGCTGGA 1379
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||
459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476
   |||||||:::|||||:::|||||:::|||||:::|||||:::|||||

```



```

310 LysLeuSerThrPlyLeuSerSerIleThrLysSerGluLysGlyTyr 326
|||||
910 AAGTTGCTTGGAAGAGCTCAGTATCAGTAAGCTGAGAGCGGAGGATA 959
|||||
326 rHsLeuThrThrGluThrProGluGlyValAlaSerLeuGlnSerArgS 343
|||||
960 CAACTTATACATATGAGACTCAGATGATGTTAGTTCCGTCGACAGCAAAA 1009
|||||
343 erLleValMetThrValProSerThrValAlaSerAsnIleLeuArgPro 359
|||||
1010 GTGTTGTAATGAGAGGTCATCATCTGTCAGTGCAGTGTCTTGGCCCT 1059
|||||
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheThrTyrProProVa 376
|||||
1060 CTTTCGAATCTGCTCAATGCAATGCATCTCAAAATATATTACCCACAGT 1109
|||||
376 lGlyAlaValThrIleSerThrProGlnAlaIleArgAspGluArgL 393
|||||
1110 TCAGAGCATATCTATCTCGTACCCGAAAGCAATCCGAAACGAATGTT 1159
|||||
393 euValAspGlyLeuLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
|||||
1160 TGATAGATGTTGTAAGTAAAGGTTTGGGCAATTGCATCCAGCCAGCAA 1209
|||||
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
|||||
1210 GGAGTTGAAACATTAGCAATCTATCTACACTCTCTCACTTCCAAATCG 1259
|||||
426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyValAlaLysA 443
|||||
1260 CGACCGCCCGGAGAAATTTGCTGTGTAACATACATTGGCGGGCTACAA 1309
|||||
443 snProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAlaSp 459
|||||
1310 ACACCGCAATTTCTCCAAAGTGAAGGTGAGTTAGTGAAGAGAGTTGAC 1359
|||||
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
|||||
1360 AGAGATTGAGGAAATGCTAATTAAGCTAATTCAGCAGCATCCACTTAA 1409
|||||
476 lValGlyValArgValThrProGlnAlaIleProGlnPheLeuValGlyH 493
|||||
1410 ATTAGAGATTAGGCTATGGCTTACAGCCATTCTCTGATTGATTGGTTC 1459
|||||
493 iLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
|||||
1460 ACTTGTATATCTTGACAGCGGCTAATCATCTACATCTTCGGGCTTAC 1509
|||||
510 GluGlyLeuPheLeuGlyLysAsnTyrValSerGlyValAlaLeuGlyAr 526
|||||
1510 GAAGGGCTATTTTGGGTGCAATTACGTCGCTGTGAGCTTACCGTACG 1559
|||||
526 gcysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 543
|||||
1560 GTCTGTGAAGAGCGCATATGAACCCGCAATTGAGGTCAACAACCTTATGT 1609
|||||
543 eArArgTyrAlaTyrLys 548
|||||
1610 CACGGTACGCTTACAAAG 1626
|||||
seq_name: gb_pat2:AX084732

```

```

seq_documentation_block:
LOCUS AX084732 1719 bp DNA PAT 09-MAR-2001
DEFINITION Sequence 1 from Patent WO0112825.
ACCESSION AX084732
VERSION AX084732.1 GI:13274985
KEYWORDS
SOURCE
ORGANISM thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

```

```

REFERENCE
1 (bases 1 to 1719)
AUTHORS Johnson,M.A., Volrath,S.L., Helfert,P.B. and Law,M.D.
TITLE Herbicide-tolerant protoporphyrinogen oxidase
JOURNAL Patent: WO 0112825-A 1 22-FEB-2001;
FEATURES
source
location/Qualifiers
1..1719
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
31..1644
/note="Arabidopsis protox-1"
/codon_start=1
/protein_id="CAC34008.1"
/db_xref="GI:13274986"
/translation="MELSLRPPTOSLPSFSPKPNLRNVKPLRLKCSVAGGPTVGS
SKTEGGGTTITTDCTIVYGGGTSGLTADALATKHPDAAPNLITVEADRVGNITTR
EENGFMEBSPNSFQSDPMLTWVDSGLKDLVLGDPAPRFLVWNGLRVPSKLT
DLPFDMLSTGKIRAGFALGIRSPGPERSEVEFERNKARDEVEFRLIRPFGSV
YADDPKLSMKNAFGKWKLEONGSIIIGTFRKLESGYNLYETPDGLVSQSKSV
GSPRKGRLMPEALISARLSKYKLSWKLISGIRKLESGYNLYETPDGLVSQSKSV
MTVPSHVASGLRLPSESANALSKLYVPPVAVASISYKELIRTECLIDSEIKFGQ
LHPRTQVETLIGTYSLSLPPNAPRGRILLNLTIGSTNTGILSKSESELYEADRD
LRKMLINPNSIDPLKLGVRWVPQALPQFLVGHFDLIDAKSSLTSSGIEGLPLGNVY
AGVALRCYGAETVALEVNEMSRAYR"
BASE COUNT 460 a 351 c 433 g 475 t
ORIGIN
alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139
alignment_block:
US-09-508-418-2 x AX084732 ..
Align seg 1/1 to: AX084732 from: 1 to: 1719
19 serSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
|||
28 TCCATGAGATTAATCTCTCTCCGCGCAGCTCAATCGCTTCTCCGTC 77
|||||
35 rLleSerLysArgAsnSerValAsnCysAsnGlyTyr.....ArgT 49
||:||||| ||| :|||
78 GTTTCGAAGCCCAAT...CTCCGATTAATGTTTATAGCCCTCTTAGAC 124
|||||
49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
|||||
125 TCCGTTGTCAGTGCCTGGACCAACCGTCGATCTCAAAATCGAA 174
|||||
66 GlyGlyProAlaAlaGluLeu.....AspCysValIleValGlyAlaG 80
|||||
175 GCGGAGAGAGGACCAACCATCACAGAGGATTTGTGATTTGCGCGGAGG 224
|||||
80 yLleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||||
225 TATTAGTGTCTTTGCAATCGCTCAGGCGCTTGCATTAAGATCTCGATG 274
|||||
95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
|||||
275 CTGCTCGCAATTTAATGTGACCGAGGCTAAGGATCGTGTGGAGCGAAC 324
|||||
110 lLeThrThrValGluArgAspGlyTyrLeuThrPheGluGlyLysProAsn 126
||| |||
325 ATTATCACTCTGTGAAGAAATGTTTCTCTCGGGAAGAAGTCCCATAG 374
|||||
126 rPheGlnProSerAspProMetLeuThrMetAlaValAlaAspCysGlyLeu 143
|||||
375 TTTTCAACCGCTCGATCTATGCTCACTATGCTGATGATGATAGTGTGA 424
|||||
143 ysAspAspLeuValLeuGlyLysProAsnAlaProArgPheValLeuTyr 159
|||||

```


ORIGIN

alignment_scores:
Quality: 2075.00 Length: 536
Ratio: 4.235 Gaps: 4
Percent Similarity: 91.418 Percent Identity: 75.373

alignment_block:
US-09-508-418-2 x ATHPROX ..

Align seg 1/1 to: ATHPROX from: 1 to: 1614

```
22 LeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSerly 38
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
7 TTAATCTCTTCCTCCGTCGACGACATCGCTTCCGCTGTTTCGAA 56
38 sArgAsnSerValAsnCysAsnGlyTirp.....ArgThrArgCys 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 GCCCAAT...CTCCGATTAAATGTTTATAGCCCTCTTAGACTCCGTTGT 103
52 eValAlaAlaLysAspTyrThrValProSerSerAlaValAspGlyLysPro 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
104 CAGTGCCGCGTGACCAACCGTCGATCTTCAAAATCGAAGCGCGAGA 153
69 AlaAlaGluLeu.....AspCysValIleValGlyAlaGlyIleSerG 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 GGCACACCATCAGCAGGATGTGTGATGTGCGGCGAGTATTAGTGG 203
83 yLeuCySIIleAlaGlnValMetSerAlaAspTyr.....ProA 96
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
204 TCTTTTCATCGCTCAGGCGCTTGCTACTAGCATCTGATGCTGCTCGA 253
96 sNLeuMetValThrGluAlaArgAspArgAlaGlyIleAsnIleThrThr 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
254 ATTATATGTGACCGAGCGCTAGACATCGTTGGAAGCAACATTCAC 303
113 ValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGlnPr 129
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 CGTGAAAGAAATGTTTCTCTGCGAAGAAAGTCCCAATAGTTTCAAC 353
129 oSerAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAsp 146
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 GTCTGATCTATGCTACTATGATGATGATGATGATGATGATGATGAT 403
146 eValValLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGly 162
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 TGGTGTGGAGATCTCTACTGCGCAAGTTGTGTGTGGAATGGGAAA 453
163 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeu 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 TTGAGCGCCGTTCCATCGAAGCTAACAGACTTACCGTTCTTGTGAT 503
179 tSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArg 196
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
504 GAGTATGTTGGTGAAGATTAGAGCTGTTTGTGTCACCTTGCATTCAG 553
196 rOserProProGlyHisGluSerValGluGlnPheValArgArgAsn 212
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 CGTCACTCCAGCTCGTGAAGAACTGTGAGAGAGTTTCTACGGCGTAA 603
213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGly 229
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
604 CTGCGTGAAGAGTTTGGAGCGCTGATTAACCGTTTGTTCAGGTGT 653
229 IlyrValGlyAspProSerLysLeuSerMetLysAlaAlaPheGlyLys 246
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
654 TTATGCTGTGATCTCTCAAAATGAGCATGAACACAGCGTTTGGGAGG 703
246 aITrPlyLeuGluGlnThrGlyGlySerIleIleGlyGlyThrPheLys 262
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 TTTGGAAACTAGACAAATGTTGGAAGCATATAGTGTACTTTTAAAG 753
```

```
263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArg 279
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
754 GCAATTCAGGAGAGAAAACGCTCCCAAGCAGACAGACAGCCGCGCT 803
279 uProLysProLysGlyGlnThrValGlySerPheArgLysGlyLeuArg 296
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
804 GCCAATAACACAGGCGCAACAGTTGTTCTTTCAGGAAGGAGCTTCGAA 853
296 eLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysSer 312
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
854 TGTTCGCCAGAGCAATATCTCCAGATTAGGTACCAAGATTAACTGTCT 903
313 TrpLysLeuSerSerIleThrLysSerGlyLysGlyGlyTyrHisLeu 329
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 TGGAGGCTCTCAGTATACATAAGCTGAGAGCGAGAGATTAACAATTAC 953
329 rTyrGluThrProGluGlyValAlaSerLeuGlnSerArgSerIleVal 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
954 ATATGAGACTCCAGATGTTAGTTTCCGTGACAGCAGCAAAAGTGTATA 1003
346 eThrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerVal 362
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1004 TGACGGTGCCATCTCATGTGCAAGTGTCCTTGCGCCCTTTTCGAA 1053
363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProValGlyAlaVal 379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1054 TCTCTGCAAAATGCATCTCAAAACTATATACCACAGTTGAGCAGGT 1103
379 IThrIleSerTyrProGluGlnAlaIleArgAspGluArgLeuValAsp 396
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1104 ATCTATCTCGTACCCGAAAGAGCAATCCGACAGAAATGTTGATAGTG 1153
396 IyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlu 412
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1154 GTGAACATAAAGGTTTGGCAATTGCATCCAGCAGCAAGAGGTTGAA 1203
413 ThrLeuGlyThrIleTyrSerSerLeuPheProAsnArgAlaProLys 429
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1204 ACATTAGGAACATATCTACAGCTCCCTCACTTTCCAATATGCGCAC 1253
429 sGlyArgValLeuLeuAsnTyrIleGlyGlyAlaValAspProGlu 446
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1254 CGGAAGAAATTTGCTGTGAACTACATTGGCGGGCTTACAAACACCG 1303
446 IeLeuSerLysThrGluSerGlnLeuValGluValAlaAspArgAspLeu 462
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1304 TTCTGTCCAACTCGAAGCTGAGTTAGTGAAGCAGATTGACAGAGATT 1353
463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValAlaGly 479
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1354 AGGAAATGCTAATTAAAGCCTAATTCACCGCATCCCTTAATTAAG 1403
479 IArgValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAsp 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1404 TAGGCTATGGCTCAGCCATCTCAGTTTCACTGATGCTACTTTGATA 1453
496 hTrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGly 512
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1454 TCCTTGACACGCGTAATTCATCTACGCTTCGCGCTACGAAAGGCTA 1503
513 PheLeuGlyLysAsnTyrValSerGlyValAlaLeuGlyArgCysValG 529
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1504 TTTTGGGTGGCAATTACGTCGTGTGAGCTTAGCGCGGTGTGAGA 1553
529 uGlyAlaTyrGluValAlaSerGluValIleThrGlyPheLeuSerArg 546
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1554 AGGGCATATGAACCGCGATGTAGGTCAACAACCTTCATGTCAAGTAC 1603
546 IatYrLys 548
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1604 CTTACAAAG 1611
```

seq_name: gb_pat2:AX084742

399 scgIpbegIyglInleuHsPProArGThrcInclYValGutHrLeuGlyT 416
|||||
1236 GGGGTTGGTCAATTGCATCCAGTACGCAAGAGTGGAAACATTAGGAA 1285
416 hrIeTySerSerIeupheProAsnArGAlaProlysGlyArVal 432
|||||
1286 CTRATACAGCTCATCTATTCCTCCACCGAGCACCTGGAAGGGTT 1335
433 LeuLeuLeuAsnTyrlIeGlylYAlaYAsnProGluIleLeuSerly 449
|||||
1336 CTACTCTGAATTACTTGGAGAGACAACTAATGGAATTTTATCGAA 1385
449 sThrcIuSerGInleuValGluValAlaAspArgAspleuArGlySMetL 466
|||||
1386 GACGGACACTGAACCTGTGGAAACAGTTGATCGAGATTGGAGAAATCC 1435
466 euIleYsProLYsAlaGInAspProLeuValValGlyValArgValTTP 482
|||||
1436 TTTATTAACCAATATGCCAGATCCATTGTGTGGGGGTGAGACTGTGG 1485
483 ProGlnAlaIleProGInpheLeuValGlyHsLeuAspThrLeuSerTh 499
|||||
1486 CCGCAAGCATTCACACAGTTCTTAGTGGCCATCTTGATCTCTAGACT 1535
499 rAlaYsAlaAlaMetAsnAspAsnGlyLeuGlyLeuPheLeuGlyG 516
|||||
1536 TCCTTAAGCTCTCATCAATAATCTGGGTTTGAAGGGCTCTTCTTGGGG 1585
516 LYAsnTyTAlSerGlyValAlaLeuGlyArGysValGluGlyAlaTyT 532
|||||
1586 GTAATTATGTCTGTGTGTGGCTTGGAGCATGGCTGAGGAGACCAT 1635
533 GluValAlaSerGlyValThrGlyPheLeuSerArgTyTAlaTyTlys 548
|||||
1636 GAGGTAGCAGCTGAAGTAAACGATTTCACAAATAGAGTGTACAAA 1683
seq_name: gb_pat2:AX084750
seq_documentation_block:
LOCUS AX084750 1784 bp DNA PAT 09-MAR-2001
DEFINITION Sequence 19 from Patent WO0112825.
ACCESSION AX084750
VERSION AX084750.1 GI:13275003
KEYWORDS
SOURCE
ORGANISM Brassica napus
rape.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 1784)
AUTHORS Johnson,M.A., Volrath,S.L., Helfetz,P.B. and Law,M.D.
TITLE Herbicide-tolerant protoporphyrinogen oxidase
JOURNAL Patent: WO 0112825-A 19 22-FEB-2001;
SynGenea Participations AG (CH)
FEATURES
source
location/Qualifiers
1..1784
/organism="Brassica napus"
/db_xref="taxon:3708"
47..1657
/note="Ollseed rape protox-1 coding sequence"
/codon_start=1
/protein_id="CAC34016.1"
/db_xref="GI:13275004"
/translation="MDLSLRPQPLSPFSNPPRSPRYKPLNLRCGSGSVGSSST
IEGGGKTYADQVIYGGISGLCTAOALVTXKDPAAKVMTEADRVGNTIPE
EÖGLMEEGNSPQSDPMLTNVDSGLKDDLYIGDTPARFVILMNKLRPVSKLTD
LPFDLMSIGKIRAGFGAIGIRSPGREGSEVEFRNLGDEVERLLLEPCSGVY
AGDPKLSMKAPEKRWKLEBNGSIIIGAFKALQANKAPKTRDRPLPKQOTVG
SFRGLTMELEAISRAGDEKYKWSLSTYKLASGSLYETPEGIVTVOSKSYM
TVPSHVASILRLPLSDSAEAELSKLYPVAASIVAKAIRECLIDELGEGOL
HPRTOKETGITYSSSLFPRAPRGVILLINTGATNGIISKSRGELVEAVDRDL
RKMLIKSSSDPLVGVKLPQALPQRLIGHIDLVDAAKASLSSGHEGLFLGNVYA
GVALGCEVGAETATQVNDPMSRYAK"

BASE COUNT 464 a 379 c 478 g 463 t
ORIGIN
alignment_scores:
Quality: 2006.50 length: 535
Ratio: 4.189 Gaps: 4
Percent Similarity: 89.533 Percent Identity: 73.084
alignment_block:
US-09-508-418-2 x AX084750 ..
Align seg 1/1 to: AX084750 from: 1 to: 1784
21 ProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerIleSe 37
||| |||||
68 CCGCAGCCATTCCTATCG.....CCATTCCAATCAT 102
37 rLYsArgAsnSerValAsnCysAsnGlyTTPArgThrArgCysSerVal 54
||| |||
103 TCCTCCG...TCGGCTCCCTCAAGCCTTCMACCTCCGTTGCTCCGTAT 149
54 lAlYsAspTyTThrValProSerSerAlaValAspGlyGlyProAlaAla 70
: : : : :
150 CCGGTGATCCGTCGCGCTCTTCTACATCGAAGCGGAGAGAGGT 199
71 Glu.....LeuAspCysValIleValGlyAlaGlyIleSerGlyLe 84
: : : : :
200 AAACCCGACAGCGGAGCTGCGTATCTCGCGGAGAGATCAGCGGCT 249
84 uCysIleAlaGInValMetSerAlaAsnTyTPro.....AsnL 97
: : : : :
250 GTGCATTCGCAAGGCGCTCGACGAAGCACCCAGACGTCGAAGAATG 299
97 eUmeTValThrcIuAlaArgAspArgAlaGlyGlyAsnIleThrThVal 113
: : : : :
300 TGATGTGACGAGGCGGAAGGACCGTGTGGAGGAATATCATCAGCGCA 349
114 GluArGAspGlyTyTLeuTrpGluGluGlyProAsnSerPheGlnProSe 130
: : : : :
350 GAGAGCAAGGTTTCTATGGGAAGAAAGTCCCAATAGCTTTCAGCGGTC 399
130 rAspProMetLeuThrMetAlaValAspCysGlyLeuLYsAspLeuV 147
|||||
400 TGATCTATGCTCACTACTATGCTGATAGTAGTGTTTGAAGATGATCTAG 449
147 AlLeuGlyAspProAsnAlaProArgPheValIleuTrpLYsGlyLeu 163
|||||
450 TCTTGGGAGATCTACTCTCCGAGTTTGTGTGTGAAGGAGACCTG 499
164 ArgProValProSerLYsLeuThrAspLeuProPheAspLeuMetSe 180
500 AGCGCGGTTCCGTCGAAGCTTACGACTGCTTCTTTCATGATGAGAG 549
180 rIleProGlyLYsLeuArgAlaGlyPheGlyProIleGlyLeuArgProS 197
||| |||||
550 TATTGAGAGGAAGATTAGACTGGGTTTGGCCATTGCTATCGACCTT 599
197 erProProGlyHsIGluGluSerValGluGlnPheValArgArgAsnLeu 213
600 CACCTCCGGGTGCTAGGAATCAGTGAAGATTGTGAAGCGCAATCTT 649
214 GlyGlyGluValAlPheGluArgLeuIleGluProPheCysSerGlyValY 230
||| |||||
650 GGTGATGAGGTTTTCAGCGCTTATTCACACCTTTTGCTAGTGTTTA 699
230 rValGlyAspProSerLYsLeuSerMetLYsAlaAlaPheGlyLYsValT 247
: : : : :
700 TGCGGAGATCCTCGGAACAGATGATGAAGACAGCTTTGGAGAGGTTT 749
247 rPLysLeuGluGluThrcIuGlySerIleIleGlyLYsIleThrPheLYsAla 263
750 GGAAGCTAGAGAGATGTGGAGACATCTGTTGTTTAAAGCA 799

| | | |
|------|---|------|
| 264 | 11elygLuiaRSeSerSerthrrProlysaLaPrCoAngsPrProArgLeuPr | 280 |
| 265 | | |
| 800 | ATTCAACGGAAAAATAAACCTCCCAAGAACCCAGATCCGGCTCTCC | 849 |
| 280 | oLysPrProLySGlGlnThrValGlySerPheArgLySGlyLeuAmeTL | 297 |
| 281 | | |
| 850 | AAAGCCAAAGGGCCAAACTGTTGGTCTTTCAGGAAGAGACTCACAAATGC | 899 |
| 297 | eupProAspAlaIleSerAlaArgLeuGlySerLyLeuLyLeuSerTrp | 313 |
| 298 | | |
| 900 | TCCCAAGCGCAATCCCGCAAGGTTGGGTGACAAAGTGAAGTTTCTTGG | 949 |
| 314 | LyLeuSerSerLleThrLySerGluLySGlyGlyTrpHisLeuThrTy | 330 |
| 315 | | |
| 950 | AAGCTTCAGATACACTAAGCTGGCCAGCGAGAAATAGCTTAACCTA | 999 |
| 330 | reluthrProGluGlyValValSerLeuGlnSerArgSerLleValMet | 347 |
| 331 | | |
| 1000 | CGAAATCCGGAGGGATAGTACCTGACGTACAGAGCAAAAGTAGTAGAGA | 1049 |
| 347 | hValProSerTrpValAlaSerAsnIleLeuArgProLeuSerValAla | 363 |
| 348 | | |
| 1050 | CTGTGCATCTCATGTGGTACTAGTCTTTCGGCCCTCTCTATATCT | 1099 |
| 364 | AlaAlaAspAlaLeuSerAsnDheTrpTrpProProValGlyAlaValTh | 380 |
| 365 | | |
| 1100 | GGAGCGAAGCCCTCCAAACCTACATACCGCGAGTTGCAGCCGATAC | 1149 |
| 380 | rLleSerTrpProGlnGlyAlaIleArgAspLyrArgLeuValAspGly | 397 |
| 381 | | |
| 1150 | CATCTCATACGGGAAGAACGACATCCGAAGCCAAAGCTTAATAGTGGTG | 1199 |
| 397 | LyLeuLySGlyPheGlyGlnLeuHisProArgTrpGlnGlyValGlyLuthr | 413 |
| 398 | | |
| 1200 | AACCTAAAGGGTTCCGGCCAGTTGCATCCACGACACCAAAAGTGGAACCT | 1249 |
| 414 | LeuGlyLuthrLleTySerSerSerLeuPheProAsnArgAlaProLySGl | 430 |
| 415 | | |
| 1250 | CTTGGAAACATATACAGTTCACTGGCTTTTCCCAACGACGACCCGCTGG | 1299 |
| 430 | yArgValLeuLeuLeuAsnTrpLleGlyGlyAlaLyAsnProGluLleL | 447 |
| 431 | | |
| 1300 | AAGAGATATGCTATTAAGACTACATCGGTGGAGACTACCAACACTGGGATCT | 1349 |
| 447 | eusErLyThrGluSerGlnLeuValGluValValAspArgAspLeuArg | 463 |
| 448 | | |
| 1350 | TATCAAAAGTCGGAAAGGTGAGTTAGTGAAGACGTAGTAGACTTGAGG | 1399 |
| 464 | LysMetLeuLleLyProLySaLaGlnAspProLeuValValGlyValArg | 480 |
| 465 | | |
| 1400 | AGAGTGTGATTAACCAAGGCTCGAACCCATCCACTTGACTTGGAGTAA | 1449 |
| 480 | gValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAspTrpL | 497 |
| 481 | | |
| 1450 | ATTATGGCCTCAAGCCATTCCTCACTTCTGTAGGTGCACATGTATTGG | 1499 |
| 497 | eusErThrAlaLySaLaIleMetAsnAspAsnGlyLeuGlyGlyLeuPhe | 513 |
| 498 | | |
| 1500 | TAGAGCGACCGAAACATACGCTCTGTCATCTGTGCATCAGAGGGCTTATTC | 1549 |
| 514 | LeuGlyGlyAsnTrpValSerGlyValAlaLeuGlyArgCysValGlnG | 530 |
| 515 | | |
| 1550 | TTGGGTGCAAAATACGTTGCCGGGTACCATTTGGGTGGTGTGGGAAGG | 1599 |
| 530 | yAlaTrpGluValAlaSerGluValThrGlyPheLeuSerArgTrpAlaTr | 547 |
| 531 | | |
| 1600 | TGCTTATGAAATCCCAACCCAAAGTAATGATTTTCATGCAAGGTATGCTT | 1649 |
| 547 | yTrpLyS 548 | |
| 1650 | | |
| 1651 | ACAAAG 1654 | |

```

seq_name: gb_p12:AF160961
seq_documentation block:
LOCUS       AF160961                1995 bp      mRNA                01-OCT-1999
DEFINITION  Cichorium intybus protoporphyrinogen IX oxidase (PPX1) mRNA.
ACCESSION   AF160961
VERSION     AF160961.1
KEYWORDS    complete cds; nuclear gene for plastid product.
SOURCE      AF160961.1  GI:6002911
  ORGANISM  Chloory.
            Cichorium intybus
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
            II; Asterales; Asteraceae; Lactuceae; Cichorium.
REFERENCE   1  (bases 1 to 1925)
  AUTHORS   Adomat,C. and Boeger,P.
  TITLE     Cloning, Sequence, Expression and Characterization of
            Protoporphyrinogen IX Oxidase from Chloory
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1925)
  AUTHORS   Adomat,C. and Boeger,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-1999) Plant Biochemistry, University of Konstanz,
            Konstanz /8434, Germany
FEATURES             Location/Qualifiers
     source          1..1925
                     /organism="Cichorium intybus"
                     /db_xref="taxon:13427"
                     /note="variety: foliosum"
     gene            1..1925
                     /gene="PPX1"
     CDS             91..1758
                     /EC_number="1.3.3.4"
                     /codon_start=1
                     /product="protoporphyrinogen IX oxidase"
                     /protein_id="AAF0194.1"
                     /db_xref="GI:6002912"
     translation     "MISLTIDCSLNCRRSSSLPPVSGSLTSKNRPLYITSPAHR
                     /KCNMRFRCSIAKDSPTITPISNENFNSQPLIDCVIVAGISGLCIAQALTKHASP
                     DVITFARDBRVGNIITFVERDGYLMEEGPNSFQPSDMLIMVYDSGLKDLVLGDPTA
                     PRFLVAMGDLKPVPSKPADLPEFLMFGPLRAGFGALGFRRSPPRESEVEFVR
                     NLGVEVERLIEPRCSGVAGADPEKLSMKAAFGKVMNLKONGSIVGAKRILDRKN
                     SOKRPRRLPKPKGQTVGSFRKQALPNAISIRLGRASRYKLCKLKSILKRN
                     LITETPGESLSQTKITIVMVPSTVASDILRPLSGLAADLSKRTYTPVAVASTYK
                     DAIADRLIDGQLKFGQLHPRSGVSTLGTITVSSLFPNRAPGAVRLNNTIGGAVN
                     PEILSTGEIIVDAVDRDLRTMLIRDAEDPLTGVVMPRAIPQFLIGHYDILDSAK
                     AALSSGGFQCMFLCGNIVSGVALGKCYEAADVAAEVMNLSQGVK"
BASE COUNT        507 a                398 c                496 g                524 t
ORIGIN
alignment_scores:
  quality: 2004.50      length: 536
  ratio: 4.194          gaps: 3
  percent similarity: 89.179      percent identity: 71.455
Alignment_block:
US-09-508-418-2 x AF160961 ..
Align seg 1/1 to: AF160961 from: 1 to: 1925
18 SerSerProLeuAlaIheLeuSnArGThrSerPheIleProPheSe 34
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
160 TCTGGTGGTCTGTGACGTCAAGAATCTAGTACTTATACAGTATAG 209
34 rSerIleSerLysArgAsnSerValAsnCySasnGlyTParGThrArgC 51
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||
210 TCCGGCGCATCCGCAAA.....TGCATATAGTGAGAGCTTCGGCT 247
51 ySerValAlaLysAspTyThrVal...ProSerSerAlaValAspGly 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||
248 GCTCATATGCGCAAGATTCGCCCAATTACTCCCTCCCATTTCAAAATGATTC 297

```

```

67 GlyProAlaIaGluLeuAspCysValIleValIleGlyIleSerG1 83
....
298 AACCTCACACCATTTGTCATGTCATTTGGGCGCCGTCATTAGCGG 347
83 YLeuCysIleAlaIleValMetSerAlaAsnTyr.....ProA 96
|||||
348 CCTTTCATTTGCCAGGCCCTTAGCCATAACAGCCCTCCGTCCTCGG 397
96 subLeuMetValThrGluAlaIleArgAspArgAlaGlyGlyAsnIleThrThr 112
.....
398 ATGTGATCGTCACCGAGCAGCAGACAGAGTGGGGGTAAATATACACG 447
113 ValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSerPheGluPr 129
|||||
448 GTTGAAGGGATGGCTATCTGGGAAGAAGCTCTTAACGCTTCAGCC 497
129 oSerAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspAspL 146
|||||
498 ATCTGATGCCATGCTCACCATGGTGGATAGTGGTTGAAGGATGAT 547
146 euValLeuGlyAspProAsnAlaProArgPheValLeuTrpLysGlyLys 162
|||||
548 TGGGTAGAGGTGACCCACAGACGCCCTTGTATATATGAGGAGATGAT 597
548 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeu 179
|||||
598 TTGAACCGGTTCTCTCCAAACCGGCTGACCTCTTCTTTGACCTCAT 647
179 tSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgP 196
|||||
648 GAGCTTTCGGAAACTCAGACCGGTTTGGTCTTGGATTCGCTCCGTC 697
196 roSerProProGluHisGluSerValGluIlePheValArgAspAsn 212
|||||
698 CTTACCTCCAGATCGCAAGAAATCGGTGAGAGTTTGTACCTAAT 747
213 LeuGlyGlyGluValPheGluArgLeuIleGluProPheCysSerGlyVa 229
|||||
748 CTTGGAGATGAGATTTTGAACGCTTGATGAACTTTTGTCTCAGGTGT 797
229 LTyValGlyAspProSerLysLeuSerMetLysAlaIlePheGlyLysV 246
|||||
798 TTATGCTGTCATCCATCAAACTTACTATGAAGCAGCATTTGGGAAGG 847
246 aLTyPlyLeuGluGluThrGlyGlySerIleIleGlyGlyThrPheLys 262
|||||
848 TCTGGAATCTGAGCAAAATGGTGTAGCATTTGTGTCGACCTTCAG 897
263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArgL 279
|||||
898 GCTATTTCAGACAGAAAGATATGCAAAAGCTCCACGAGCCGAGGTT 947
279 uProLysProLysGlyGlnThrValIleGlySerPheArgLysGlyLeuArgM 296
|||||
948 ACCGAAACCAAGGGCCAACTGTGTGATCTTTAGAGAGGACCAAGCGA 997
296 eTLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysSer 312
|||||
998 TGTTCCTTAATGCAATCTCAACGAGTTAGTACACAGTGAATTTGCT 1047
313 TrpLysLeuSerSerIleThrLysSerGluLysGlyTyrHisIleuTh 329
|||||
1048 TCGAAGCTCAGAGATATTCAAAATTTGGAATAGAGATTAATATTTGAC 1097
329 rTyTrpGluThrProGluGlyValValSerLeuGlnSerArgSerIleValM 346
|||||
1098 ATATGAAGAACCCACAGAGATTTGAAGTCTCCACACATRAAACTTCGTGA 1147
346 eThrValProSerTyValAlaSerAsnIleLeuArgProLeuSerVal 362
|||||
1148 TGACGTGTCATCATCGTGGGAGTACTGTTGTCGTCGCTTTCGTTG 1197

```

```

363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaVa 379
|||||
1198 GGTGACGACGATGTCATTTGCAAAATTTATATCTCCGCTTGACGCTGT 1247
379 LThrIleSerTyrProGlnGluAlaIleArgAspGluTrpGluValAspG 396
|||||
1248 ATCAATTTATATCCAAAGACGCAATTCGTCTGACCGGCTGATTTAG 1297
396 LysIleuLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlu 412
|||||
1298 GTCAACTCAAGCTTTTGGGCAATTGACATCCACAGAAATCAAGGGTGA 1347
413 ThrLeuGlyThrIleTyrSerSerLeuPheProAsnArgAlaProLy 429
|||||
1348 ACTTAGGTACGATCTACAGTTCAATCTCTTCCCTAACCGAGCCGACC 1397
429 sGlyArgValIleLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGlu 446
|||||
1398 TGGAGGGTCTGCTTGAACATCGAGGAGGCTACAAATCTGAAA 1447
446 LeuSerLysThrGluSerGlnLeuValGluValAlaAspArgAspLeu 462
|||||
1448 TTCTATCAAAAGACGAGGGCAATGTGATGCGGTGACCGGACCTTA 1497
463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValGlyVa 479
|||||
1498 CGGACGATGCTGATTAAGCCGTATGCGGAAGATCCATTGACTGGGGGT 1547
479 LArgValTrpProGlnAlaIleProGlnPheLeuValGlyHisLeuAsp 496
|||||
1548 GCGGGTGTGCTCGACACATCCCGCAGTTTGTGATCGGTCAATTAGCA 1597
496 hLeuSerThrAlaLysAlaIleMetAsnAspAsnGlyLeuGluGlyLeu 512
|||||
1598 TTTCAGATCTCGCAAAAGCTCTGTGAGTACGCGTGATTCACAGGTATG 1647
513 PheLeuGlyGlyAsnTyValSerGlyValAlaLeuGlyArgCysValG1 529
|||||
1648 TTCTTGTGTGACACTTGTGTCTGTGTGCTTGTAGTAATGTGTGCA 1697
529 uGlyAlaTyGluValAlaSerGluValThrGlyPheLeuSerArgTyra 546
|||||
1698 GCGCGCTTATGATGTTCCCGCTGAGTAATGAACTTTGTGTGCAAGGGG 1747
546 LAtyTyLys 548
:::|||||
1748 TGTACAAG 1755

seq_name: gb_p12:AF218052

seq_documentation_block: 1608 bp mRNA PLN 15-FEB-2000
LOCUS AF218052
DEFINITION Zea mays protoporphyrinogen IX oxidase mRNA, complete cds.
ACCESSION AF218052
VERSION AF218052.1 GI:6715440
KEYWORDS
SOURCE
ORGANISM Zea mays.
Zea mays.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1608)
REFERENCE
1 Volrath, S.L.
Direct Submission
Submitted (20-DEC-1999) Novartis Agricultural Biotechnology, 3054
Cornwallis Road, Research Triangle Park, NC 27709, USA
FEATURES
source
1..1608
/organism="Zea mays"
/cultivar="B73 inbred"
/db_xref="taxon:4577"
1..1608
/codon_start=1
CDS

```

```

/product="protoporphyrinogen IX oxidase"
/protein_id="AAP6417.1"
/db_xref="GI:6715441"
/translation="MAAATATAMATAAPSLPLNGTRIPARLRHRLSGVCAAVAGCAAE
AASAGARLSADCVYGVGGISGLCTAOALTRRGVGVLTETARAPGGNTTVEREE
EYLLIEEGNSFOPSDPVLPMADVPSGLKDLVFGDENARFELMGKRLRPVSPADP
PEFDMSIPGRLRAGLALGIRPPPEKRESVVEFRRNLGAVERLRIEFGCSGYA
GPPSLSMKAAKPAKGVMLKEENGSLTGITGTIOEKSKMPKPRARLRPKGOTVAS
PRKGLAMPNAITSSLSGSKYLRSLKRLSTIKSDSKYVLEYETTPGCVSYQAKSVYTH
IPSYASNTLRPLSSDADADLSRFTYPPVAAYVATSPKAIKBECLDIGELGEGQLH
PSSQSEVLTGLTYSSSLFPPNAPAGRYLLNTIGATNGTIVYSKESLVAADVDRDL
KMLINSTAVDLVLGVYVMPQALPQFLVGLHDLLEAKAALDRCGYDGLFLFGANYVAG
VALGRCVGAESQSLDSFLTKAYK"
BASE COUNT      360 a      430 c      464 g      354 t
ORIGIN

```

```
alignment_scores:
  quality: 1989.50      length: 545
  ratio: 4.260          gaps: 6
Percent Similarity: 85.688  Percent Identity: 72.294
```

US-09-508-418-2 X AF218052

Align seg 1/1 to: AF218052 from: 1 to: 1608

17 serSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProph 33
 31 ACCGCTGCATGCCCTTA.....CTCAAGCGGAGCCGA..ATACCTGC 71
 33 eSerSerIleSerIysArgAsnSerValAsnCysAsnGlyTTPargThra 50
 72 GCGGCTCCGCCATCCAGGACTACAGCTGTCGCTGC..... 105
 50 rGcYSerValAlaLysAspYrThrValProSerSerAlaValAspGly 66
 106GCTGCTGTGCGCGC 120
 67 GlyProAla.....AlaGluLeuAspCy 74
 121 GCGCGGCGCGAGCCACCGCATCCACGCGCGCGGCTGTCCGCGGACTG 170
 74 sValIleValGlyAlaGlyIleSerGlyLeuCysIleAlaGlnValMet 91
 171 GCTGTGTGTGGCGGAGGCGATCAGTGGCTCTGACCGCGCGAGGCGCTGG 220
 91 eRAlaAsnTyr.....ProAsnLeuMetValThrGluAlaArgAspArg 105
 221 CCAGCGCGCACGCGCTCGGGAGCGCTGTTCACGAGAGCGCCGCGCCGC 270
 106 AlaGlyGlyAsnIleThrThrValGluArg.....AspGlyYrIleutr 120
 271 CCGCGCGGCAATTCACCGCTGAGCGCGCCGAGGAAGGATACCTCTG 320
 120 pGluGluGlyProAsnSerPheGlnProSerAspProMetLeuThrMet 137
 321 GGAGAGGGGTCCCAACACTTCACGCCCTCGACGCCGCTTCTCACCAATGG 370
 137 lAvAlaSPcysGlyLeuLysAspAspLeuValLeuGlyAspProAsnAla 153
 371 CCGTGGACAGCGGACATGAAGATATGCTGTTTGGGGAGCCAAACGCG 420
 154 ProArgPheValLeuTrpLysGlyLysLeuArgProValProSerLysLe 170
 421 CGCGCTTTCGTCTGTGGAGGAGAGCTGAGGCCCTGCGCATCCAAAGCC 470
 170 uThrAspLeuProPhePheAspLeuMetSerIleProGlyLysLeuArg 187
 471 CGCGACACTCCGCTTTCGATCTCATGAGAGATCCAGGAGAGCTCAGGG 520
 187 lAgLyPheGlyProIleGlyLeuArgProSerProProGlyHHisGluGlu 203

| | | |
|------|---|------|
| 521 | CCGGTCATGAGGCGCCTTGAGCAATCCGCGCCGCTCCACAGCGGAGAG | 570 |
| 204 | SeRTVIGLUGLInPhvAlArArqsmLseuGLyglYvAlPheGLAr | 220 |
| 571 | TCAgTGGAGAGATTCGTGCGCGCAACTCGGTGAGAGCTTTGAGCG | 620 |
| 220 | gLeuIleGLuPrOphEcYsSerGLYvAlYvAlGLYAsPProSerLYSL | 237 |
| 621 | CCCTCATTTGAGCCTTCTCTGCTCAGGGGTGTAGCGGTGATCTTCTAAGC | 670 |
| 237 | eUSeMeLtySLAlaLapheGLYvSLrTPlySLeugLIGLInrGLY | 253 |
| 671 | TCAcATGAGAGCTGCATTTGGGAAGGTTTGCGGTTTGGAGAAACTGCA | 720 |
| 254 | gLYSerLleIleGLYGLYThrPheLYSLAlleLYSGluArqSerSerTh | 270 |
| 721 | GGTAACTATTATTTGGGGAACCATCAAGCAATTCAGAGAGACGACAAAGA | 770 |
| 270 | rProLYSLArProArqAsPProArLeuProLYSProLYSGLYInrV | 287 |
| 771 | TCCAAACACCCGAGAGATGCCCGCTTCCAGACCAAAAGGACGACAG | 820 |
| 287 | aGLYSerPheArqLYSGLYLeuArqMeLLeuProAsPaLleSerAla | 303 |
| 821 | TTGCATCTTTTCAGGAGAGGCTCTTGCCAGCTTCCAAAAGCCATTAACTCC | 870 |
| 304 | ArqLeuGLYSerLYSLeUlySeuSerTrPlyLYSeuSerSerLleThLY | 320 |
| 871 | ACGTTGGGTAGTAAAGCTCAACTATCATAGAAACCTCAGACACTTACAA | 920 |
| 320 | sSerGLYSGLYGLYrThSLeUthrTYrGLInrProGLUGLYvAlY | 337 |
| 921 | ATCAGATGCAAGGAGATATGTTTGGAGTATGAAAGCCAAAGGCGTGC | 970 |
| 337 | aLSeuLeuInSerArqSerLleVAlMeLThrVAlProSerTYrVAlAla | 353 |
| 971 | TTTCGGTGAGGCTTAAGTGTATCATGACATTCATCATATGTGTCT | 1020 |
| 354 | SeRAsuLleUArqProLeuSerVAlAlaAlaAsPaLaleuSerAs | 370 |
| 1021 | ACCACACTTTTCCGTCACCTTTTCAAGCAGCATGCGAGATGCTTACTACAG | 1070 |
| 370 | nPheYrTYrProPArVAlGLYvAlVAlrLleSerYrProGLInGLuA | 387 |
| 1071 | ATTCTATATTACACGGGTGCGTGTAACTGTTTCGTATCCAAAGAGAG | 1120 |
| 387 | lAlleArqAsPaLArqLeuVAlAsPGLYGLInLeuLYSGLYPheGLYIn | 403 |
| 1121 | CAATTAGAAAGAAATCTTAATTTGATGGGAACTCCAGGCTTTGGCCAG | 1170 |
| 404 | LeuHLSProArqTrhGLInGLYvAlGLInrLeuGLYTrLleYrSerSe | 420 |
| 1171 | TTCGATCCACGTAAGCAAGGAGTTGAGACATTAGAACAATATACAGTTC | 1220 |
| 420 | rSerLeuPhePrOAsnArqVAlArProLYSGLYArqYAlleuLeuAsnT | 437 |
| 1221 | CTTCACCTTTTCCAAATCGTCTCTGACGGTAAAGGTTGTACTCTTAACCT | 1270 |
| 437 | YrLleGLYGLYAlaLYSLAsnProGLInLleSeuSerLYSThrGLInSerGLn | 453 |
| 1271 | ACATAGGAGAGCTCTCAAAACACAGGAATTTGTTCCAAAGCTGAAGGTAG | 1320 |
| 454 | LeuVAlGLYvAlVAlAsPArqAsPLeuArqLYSMeLLeuLleLYSProLY | 470 |
| 1321 | CTGGTCGGAAGCGTGTACCGCTGACCTCCGAAAAAGCTTAAATCTTAC | 1370 |
| 470 | sAlaLYSLAsPProLeuVAlVAlGLYvAlArqVAlrProGLInAlleP | 487 |
| 1371 | ACCAAGTGAACCTTAGTCCTTGCGGTTCGAGTTTGCCACACAGCCATAC | 1420 |
| 487 | rGLInrPheUvAlGLYInSLeUAsPTrLleUSeTrThrAlaLYSLAla | 503 |
| 1421 | CTCAGTCTCTGTATGACATCTTGATCTTTGGGAAGCCGCAAAAGCTGCC | 1470 |

```

504 MetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyGlyAsnTyrValSerG 520
      ::::: ::::: ::::: ::::: ::::: :::::
1471 CTGGACCGAGGTGGCTACGATGGGCTGTCTCTAGGAGGAGCACTATGTTC 1520
      ::::: ::::: ::::: ::::: ::::: :::::
520 rglValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAlaSerG 537
      ::::: ::::: ::::: ::::: ::::: :::::
1521 AGGAGTTCCTGGCGATGCTGTGAGGCGCGGTATGAAGTGCCTCGC 1570
      ::::: ::::: ::::: ::::: ::::: :::::
537 LuValThrGlyPheLeuSerArgTyrAlaTyrLys 548
      ::::: ::::: ::::: ::::: ::::: :::::
1571 AATATCTGACTCTCTGACCAATATGCTACAG 1605
      ::::: ::::: ::::: ::::: ::::: :::::
seq_name: gb_pat2:AX084740

seq_documentation_block:
LOCUS      AX084740      1811 bp      DNA
DEFINITION Sequence 9 from Patent WO0112825.
ACCESSION  AX084740
VERSION     AX084740.1  GI:13274993
KEYWORDS
SOURCE
  ORGANISM
    bread wheat.
    Triticum aestivum
    Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
    Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;
    Triticum.
    1 (bases 1 to 1811)
REFERENCE  1
  AUTHORS  Johnson,M.A., Volrath,S.L., Heifetz,P.B. and Law,M.D.
  TITLE    Herbicide-tolerant protoporphyrinogen oxidase
  JOURNAL  Patent: WO 0112825-A 9 22-FEB-2001;
            Syngenta Participations AG (CH)
FEATURES
  source
    1..1811
    /organism="Triticum aestivum"
    /db_xref="taxon:4565"
    <3..1589
    /note="wheat Protop-1"
    /codon_start=-1
    /protein_id="CAC34012.1"
    /db_xref="GI:13274994"
    /translation="ATMATATVAASPLRGRTVGRPHRVPRCATASATETPAAPCV
    RLSECVIVAGISGLCTAOLATRYGVSLDVTLEARDRGNTTYERDDEGLME
    GPNSEFQSDPVLTMVAVDSGLKDLVFGDPNAPRVVLEGLRPVPSKPGDLPPSLMS
    IPGRLRAGLGLATGIRPPPRESEVEEVEFRNRNLAQVEFLIEFCGVTAGDPSKLS
    MKPAKGVWRIEETIGSITIGTITKADDKRNKPPDPPLPAKGVTVASFKGLAM
    LPAATASRLGSKVLSWKLSTIKADNQGVLTGPEGLVSVQAKSVIMTIFRSYAS
    DILRLPSIDADALSKFYPPVAATVSYKFAIRKCLDGLDGLGGLHPRSGQVE
    TLGTYSSSLFPPNAPAGARVLLNVTIGSTNTGIVSTESDLVGAVDRLRKMLNPR
    AADPLAGVRVMPQAIPOFLIGHDLRLAAKSAKSLGCGGVDGLFLGNGYVAGVALGRCI
    EGAVESAQVSDPFTKYAK"
BASE COUNT  423 a      463 c      517 g      408 t
ORIGIN
alignment_scores:
  Quality: 1987.00      Length: 505
  Ratio: 4.310      Gaps: 3
  Percent Similarity: 91.287      Percent Identity: 75.050
alignment_block:
US-09-508-418-2 x AX084740 ..
Align seg 1/1 to: AX084740 from: 1 to: 1811
48 ArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAlaVala 64
   ||| ||||| ::::: ::::: ::::: :::::
78 CCGCGCGCTGGCGTACCGGACGCGGACGACGACGACGCGCGCGGCG... 125
   ||| ||||| ::::: ::::: ::::: :::::
64 LaspGlyGlyProAlaAlaGluLeuAspCysValAlaValGlyAlaGlyT 81
   ||| ||||| ::::: ::::: ::::: :::::
126 ....CCCGGCGTGGCGGCTGCGCGGCAATGCGTCAFTGTGGCGCGCGGCA 171
   ||| ||||| ::::: ::::: ::::: :::::
81 LserGlyLeuGlyLeuGluValMetSerAlaAsnTyr.....Pro 95

```

```

172 TCAGGCGGCTCGACCGCGAGGGCGTGGCCACCGCATGAGCGGTAGC 221
      ||||| ::::: ::::: ::::: :::::
96 AsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsnIleThrTh 112
      ::::: ::::: ::::: ::::: ::::: :::::
222 GACCTGCTGCTACGAGAGCCCGACCGCCGGCGCAACATACAC 271
      ||||| ::::: ::::: ::::: :::::
112 rValGluArg.....AspGlyTyrLeuThrPgluGluGlyProAsnSerP 127
      ||||| ::::: ::::: ::::: :::::
272 CGTCAGGCTGCCCGACGAGGGGTACCTGTGGAGAGGAGGCCAACAC 321
      ||||| ::::: ::::: ::::: :::::
127 heGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuLys 143
      ||||| ::::: ::::: ::::: :::::
322 TCCACCGCTCCGACCGCGTCCCTCACCATGCGCGTGGACAGCGGCTCA 371
      ||||| ::::: ::::: ::::: :::::
144 AspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTyrPly 160
      ||||| ::::: ::::: ::::: :::::
372 GATGACTGGTGTGGGGAGCCCAACCGCGGCTTCGTGCTGTGGGA 421
      ||||| ::::: ::::: ::::: :::::
160 sGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePhe 177
      ||||| ::::: ::::: ::::: :::::
422 GGGAGAGCTGAGCGCGGTGCGTGAAGCCAGCGACCTGCTTCTTCA 471
      ||||| ::::: ::::: ::::: :::::
177 sPLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleGly 193
      ||||| ::::: ::::: ::::: :::::
472 GCCTCATGATGATCCCTGGGAAGCTCAGCGCGGCTTGGCGCCTCGGC 521
      ||||| ::::: ::::: ::::: :::::
194 LeuArgProSerProProGlyHisGluGluSerValGluGlnPheValAr 210
      ||||| ::::: ::::: ::::: :::::
522 ATGCGCCACCTCCTCCAGCGCGGAGAGAGTGGTGAAGAGTTTGTGCG 571
      ||||| ::::: ::::: ::::: :::::
210 gArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCys 227
      ||||| ::::: ::::: ::::: :::::
572 CCGCAACCTCGGTGCGGAGTCTTGAAGCGCTCATGAGCCTTCTGCT 621
      ||||| ::::: ::::: ::::: :::::
227 ergLysValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaPhe 243
      ||||| ::::: ::::: ::::: :::::
622 CAGGTATATGCTGGTGTATCTTGAAGCTTATATGAAGCGTCAATT 671
      ||||| ::::: ::::: ::::: :::::
244 GlyLysValTyrPlyLysLeuGluGluThrGlyLysSerIleIleGlyGly 260
      ||||| ::::: ::::: ::::: :::::
672 GGGAGAGCTGAGAGTGGAGAGATGGAGTGTATATGTTGTTGAC 721
      ||||| ::::: ::::: ::::: :::::
260 rPheLysAlaIleLysGluArgSerSerThrProLysAlaAlaProArgasp 277
      ||||| ::::: ::::: ::::: :::::
722 CATCAAGCGCATTCAGCATTAAGGGAAGAACCCCAACCGCGCAAGGATC 771
      ||||| ::::: ::::: ::::: :::::
277 rArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysGly 293
      ||||| ::::: ::::: ::::: :::::
772 CCGGACTTCCGCAACAAAGGAGACAGCGTGGCATCTTTCAGGAAGGT 821
      ||||| ::::: ::::: ::::: :::::
294 LeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 310
      ||||| ::::: ::::: ::::: :::::
822 CTAGCATGCTCCCGAATGCCATGCACTAGCGTGGGTAGTAAGCA 871
      ||||| ::::: ::::: ::::: :::::
310 sLeuSerTyrLysLeuSerSerIleThrLysSerGluLysGlyGlyTyr 327
      ||||| ::::: ::::: ::::: :::::
872 GCTGTCATGAGAGCTTACGACCATTAAGGCGGACCAACCAAGATATG 921
      ||||| ::::: ::::: ::::: :::::
327 sLeuThrTyrGluThrProGluGlyValAlaSerLeuGlnSerArgSer 343
      ||||| ::::: ::::: ::::: :::::
922 TATTAGTTATGAACACACAGAGGACTTGTTCAGTGCAGGCTAAAGT 971
      ||||| ::::: ::::: ::::: :::::
344 IleValMetThrValProSerTyrValAlaSerAsnIleLeuArgProLe 360
      ||||| ::::: ::::: ::::: :::::
972 GTTATCATGACATCCGCTCATATGTTGTAGTATATCTTGGCCGACT 1021
      ||||| ::::: ::::: ::::: :::::
360 uSerValAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValG 377
      ||||| ::::: ::::: ::::: :::::
1022 TTCAATGATGACAGATGACATCTCAAAATCTATTATCCCGCAGTTG 1071
      ||||| ::::: ::::: ::::: :::::
377 LyAlaValThrIleSerTyrProGluGluAlaIleArgAspGluArgLeu 393
      ||||| ::::: ::::: ::::: :::::

```


654 GACAGTTCATCTTTCAGGAGGCTCTGCCATGCTTCCAAATGCCATT 703
302 eraIaArGLeuGlySerLyLeuLyLeuSerTrpLyLeuSerSerIle 318
704 CATTCCACTTGGGTAGTAAGTCAAACTATCATGGAACACAGAGATT 753
319 ThrLySerSerLyLeuGlyTyrHsLeuThrTyrGluThrProGlu 335
754 ACAAAATCAGATGACAGAGATATGTTTGGAGTATGAAACGCCAGAG 803
335 yAlaValSerLeuGlnSerArgSerIleValMetThrValProSerTyr 352
804 GGTGTTTCGGTGCAGGCTAAAGTGTATCATGACTATTCCATCATAG 853
352 aIaIaSerAsnIleLeuArgProLeuSerValaIaIaIaIaSerLeu 368
854 TTCTTACCAACATTTTGCCTCCTTCAAGCATGCTGACATGCTCTA 903
369 SerAsnPhetTyrTrpProValGlyAlaValThrIleSerTyrProG 385
904 TCAGATTTATATCCACCGGTGCTGCTGTAACGTGTTCCATCCAA 953
385 ngIaIaIaIaIaArgAspIuArgLeuValAspGlyIuLeuLyGlyPheG 402
954 GGAAGCAATTAGAAAAGATGCTTAATGATGGGAACTCCAGGCTTGG 1003
402 LyGlnLeuHsProArgTyrGlnGlyValGluThrLeuGlyThrIleTyr 418
1004 GCCAGTTCATCCACGATGACAGAGATTGACATTTAGAACATATATAC 1053
419 SerSerSerLeuPheProAsnArgAlaProLyGlyArgValLeuLeu 435
1054 AGTTCCTCACTCTTTCCAAATCGTCTCCGACGGTAAAGGCTTACTCT 1103
435 uAsnTyrIleGlyAlaLyAsnProGluIleLeuSerLyThrGluS 452
1104 AAACATCATGAGAGTGTCTACAAACACAGGAATTTTCCAAAGACTGAA 1153
452 erGlnLeuValGluValAlaAspArgAspLeuArgLyMetLeuIleLyS 468
1154 GTAGCTGTGCGAAGCAGTTGACCGTGCACCTCCGAAAATGCTTATTAAT 1203
469 ProLyAlaGlnAspProLeuValValGlyValArgValTrpProGln 485
1204 TCTACAGCACTGACCCCTTAGTCTGCTGTTGCGAGTTGGCCACAAAC 1253
485 aIleProGlnPheLeuValGlyHsLeuAspThrLeuSerThrAlaLySA 502
1254 CATACCTCAGTTCCTGTAGGACATCTCATCTCTGGAAGCGCGCAAG 1303
502 IaIaMetAsnAspAsnGlyLeuGlyLeuPheLeuGlyLyAsnTyr 518
1304 CTTCCTGTGACGAGGTGCTAGAGAGGCTCTCTTCCAGAGGAACCTAT 1353
519 ValSerGlyValAlaLeuGlyArgCysValGlyValaIaTyrGluVal 535
1354 GTTGCAGAGAGTTCCTGTGAGCATGCTGAGGCGCTATGAAAGTGC 1403
535 aSerGlyValaIaThrGlyPheLeuSerArgTyrAlaTyrLyS 548
1404 CTTCGCAAAATATCTGACTTCTGACCAAGTATGCTCATCAAG 1443
seq_name: gb_pat1:A78959
seq_documentation_block:
LOCUS A78959 1698 bp DNA PAT 20-OCT-1999
DEFINITION Sequence 13 from Patent WO9829554.
ACCESSION A78959
VERSION A78959.1 GI:6092087
KEYWORDS
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

REFERENCE 1 (bases 1 to 1698)
AUTHORS Ishige,F. and Sato,R.
TITLE METHODS OF CONFERRING PRO-INHIBITING HERBICIDE RESISTANCE TO PLANTS
BY GENE MANIPULATION
JOURNAL Patent: WO 9829554-A 13 09-JUL-1998;
ISHIGE FUMIHARU (JP); SATO RYO (JP)
FEATURES
source
1.1698
/organism="Zea mays"
/strain="B73 INBRED"
/db_xref="taxon:4577"
2.1453
/codon_start=1
/product="PROTOPORPHYRINOGEN OXIDASE"
/protein_id="CAB59133.1"
/db_xref="GI:6092088"
/translation="NSADCVVVGISGCTAQAALATRHGVDVLTENARPGNIT
TVRPEREGYLMDEGPNPSQSPDPLVTMAVDSGLDVLGDPNAPFVLEKRLPEP
SKPADLPFPLMELIPKLRAGLALGIRPPGREGSEVEFVRNRNGAEVFRLEP
CSGVVAGDPKSLSMKAAPGKVMRLFEETGSIIGCTIKTIOESKNPKPRDARLRPK
GRTVAAPRKLAMLPNATISSUSKYSKLSKLTSTKSDKGVLETFEPVAVYA
KSYIMTPSTVASNIIIRPLSSDADLSRFYPPVAAYVSYPKAIRRECLIDELQ
GFQDLHPRSGVETLGTIYSSLEFNRADGRVLTIGATNTGIVSKTESLEVA
VDRDLKMLINSTAVDPLVGVVRWPAIPEFLVGHLDLLEAKAALDRGVDGLFG
GNVAVGVALGRVCEGAVESASQISDFLTVAVK"
BASE COUNT 416 a 396 c 454 g 432 t
ORIGIN
alignment_scores:
Quality: 1978.00 Length: 480
Ratio: 4.455 Gaps: 2
Percent Similarity: 92.500 Percent Identity: 78.750
alignment_block:
US-09-508-418-2 x A78959 ..
Align seg 1/1 to: A78959 from: 1 to: 1698
73 AspCysValIleValGlyAlaGlyIleSerLyLeuCysIleAlaGlnVa 89
11 GACTCGTGTGTGGTGGGAGCATCATGTGCTCTGACCGCGCAGGC 60
89 IMetSerAlaAsnTyr.....ProAsnLeuMetValThrGluAlaArgA 104
61 GCTGCCACGCGCGCAGCGCTCGGGGAGCTCTTGTCAAGAGCGCGCG 110
104 sPArGAlaGlyLyAsnIleThrThrValGluArg.....AspGlyTyr 118
111 CCCGCCCGCGCGCAACATATACACCGTCGAGCGCCCCGAGGAAGGTTAC 160
119 LeuTrpGluGluGlyProAsnSerPheGlnProSerAspProMetLeuTh 135
161 CTCTGGAGGAGGGTCCCAACAGCTTCCAGCCCTCCACACCCCGTTCAC 210
135 IMetAlaValAspCysGlyLeuLyAspAspLeuValLeuGlyAspProA 152
211 CATGCCCTGACAGCGAGCTGAAGCATGACTTGTGTTTGGGACCACCA 260
152 sNaIaProArgPheValLeuTrpLySlyLyLeuArgProValProSer 168
261 ACGGCGCGCGTTCGTGCTGTGGAGGGAAGCTGAGGCGCGTGCATTC 310
169 LysLeuThrAspLeuProPhePheAspLeuMetSerIleProGlyLySle 185
311 AAGCCCGCGACCTCCGTTCTTGATCTCATGAGCATCCACAGGGAAGCT 360
185 uArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGlyHsG 202
361 CAGGCGCGGTCTAGCGCGGCTGGCATCCGCGCCCTCTCCAGGCGCGG 410

```
202 lucIuSerValIGluInPheValArgArgAsnLeuGlyGlyGluValAlaPhe 218
|||||.....:|||||:|||||.....:|||||
411 AAGAGTCAGTGAGGAGAGTCTGTCGCCGCAACCTCGTGTGAGGCTTT 460
|||||.....:|||||:|||||.....:|||||
219 GluArgLeuIleGluProPheCysSerGlyValTyrValGlyAspProse 235
|||||.....:|||||:|||||.....:|||||
461 GAGCGCTCATGTGAGCTTCTGCTCAGGTGTCTATCTGTCATCTTC 510
|||||.....:|||||:|||||.....:|||||
235 rLysLeuSerMetLysAlaAlaPheGlyLysValTyrPLysLeuGluLut 252
|||||.....:|||||:|||||.....:|||||
511 TAAGCTCAGCATGAAAGCTCATTTGGGAAGGTTGGCGGTGGAAGAAA 560
|||||.....:|||||:|||||.....:|||||
252 hrgIyGlySerIleIleGlyGlyThrPheLysAlaIleLysGluArgSer 268
|||||.....:|||||:|||||.....:|||||
561 CTGGAGGTAGATTATTGTGTGAACCATCAAGCAATTCAAGAGAGAGC 610
|||||.....:|||||:|||||.....:|||||
269 SerThrProLysAlaProArgAspProArgLeuProLysProLysGlyI 285
|||||.....:|||||:|||||.....:|||||
611 AAGAAATCCAAACCAAGGAGGATGCCGCTCCGAAAGCCAAAGGGCA 660
|||||.....:|||||:|||||.....:|||||
285 nThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAlaIleS 302
|||||.....:|||||:|||||.....:|||||
661 GACAGTTGCATCTTTCAGGAAGGCTTGGCATGCTTCCAAATGCCATTA 710
|||||.....:|||||:|||||.....:|||||
302 eraIlaArgLeuGlySerLysLeuLysLeuSerTyrPLysLeuSerSerI 318
|||||.....:|||||:|||||.....:|||||
711 CATCCAGCTTGGGTAGTAAGTCAACTATCATGGAACCTCAGAGCATTT 760
|||||.....:|||||:|||||.....:|||||
319 ThrLysSerGluLysGlyTyrHisLeuThrTyrGluThrProGluI 335
|||||.....:|||||:|||||.....:|||||
761 ACAAAATCAGATGACAAGGATATGTTTGAGTATGAAGCCAGCAAGG 810
|||||.....:|||||:|||||.....:|||||
335 yValValSerLeuInSerArgSerIleValMetThrValProSerTyrV 352
|||||.....:|||||:|||||.....:|||||
811 GGTGTTTGGTGCAAGCTAAAGTGTATCATGATCTATTCATCATATG 860
|||||.....:|||||:|||||.....:|||||
352 aAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAlaLeu 368
|||||.....:|||||:|||||.....:|||||
861 TTGCTAGCAACATTTTGCTCCACTTTCAAAGCATGCTCGAGATGCTTA 910
|||||.....:|||||:|||||.....:|||||
369 SerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrProG 385
|||||.....:|||||:|||||.....:|||||
911 TCAAGATCTATATATCCACCGTGTGCTGTAAGTTCGTATCCAAA 960
|||||.....:|||||:|||||.....:|||||
385 nGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGlyPheG 402
|||||.....:|||||:|||||.....:|||||
961 GGAAGCAATTAAGAAAGATCTTAATTGATGGGAATCCAGGGCTTTG 1010
|||||.....:|||||:|||||.....:|||||
402 lYgInLeuHisProArgThrGlnGlyValGluThrLeuGlyThrIleTyr 418
|||||.....:|||||:|||||.....:|||||
1011 GCCAGTTGCATCCACGTAGTCAGAGAGTTGAGACATTAGGAACAATATAC 1060
|||||.....:|||||:|||||.....:|||||
419 SerSerSerLeuPheProAsnArgAlaProLysGlyArgValLeuLeuLe 435
|||||.....:|||||:|||||.....:|||||
1061 AGTTCCTCAGCTTTCCAAATCGTGCTCTGACGTAGGGTGTACTTCT 1110
|||||.....:|||||:|||||.....:|||||
435 uAsnTyrIleGlyValAlaLysAsnProGluIleLeuSerLysThrGlnS 452
|||||.....:|||||:|||||.....:|||||
1111 AAACACTAGTAGAGGTGCTACAAACACAGGAATGTTTCCAGACTGAAA 1160
|||||.....:|||||:|||||.....:|||||
452 erGInLeuValGlyValValAspArgAspLeuArgLysMetLeuIleLys 468
|||||.....:|||||:|||||.....:|||||
1161 GTGAGCTGTGTGAAGCAAGTGAACGTGACCTCCGAAAAATGCTTATAAT 1210
|||||.....:|||||:|||||.....:|||||
469 ProLysAlaGlnAspProLeuValValGlyValArgValTyrProGlnAl 485
|||||.....:|||||:|||||.....:|||||
1211 TCTACAGCAGTGGAACCTTGTAGTCTGTGGTGTGAGATTGGCCACAAGC 1260
|||||.....:|||||:|||||.....:|||||
485 aIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAlaLysA 502
|||||.....:|||||:|||||.....:|||||
1261 CATACCTCAGTTCCTGGTAGGACATCTTGATCTTGTGGAAGCCGCAAAAG 1310
|||||.....:|||||:|||||.....:|||||
* 502 laAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyLysnTyr 518
|||||.....:|||||:|||||.....:|||||
|||||.....:|||||:|||||.....:|||||
1311 CTGCCCTGGACCGAGAGTGGCTACGATGGGCTGTCTCTAGGAGGAACTAT 1360
|||||.....:|||||:|||||.....:|||||
519 ValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyrGluValAl 535
|||||.....:|||||:|||||.....:|||||
1361 GTTGCAGAGAGTTGCCCTGGCGAGATGCCGTGAGGGCGGCTATGAAGAATGC 1410
|||||.....:|||||:|||||.....:|||||
535 aSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
|||||.....:|||||:|||||.....:|||||
1411 CTGCAAAATATCTGACTTCTTGAACCAAGTATGCCCTACAAG 1450
|||||.....:|||||:|||||.....:|||||
```


OM of: US-09-508-418-2 to: N_Geneseq_0601.* out_format : pfs
Date: Jul 3, 2001 11:40 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=framet_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09508418/r/unat_03072001.100017.24474/app-query.fasta_1.613
-DB=N_Geneseq_0601 -OFMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEC=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blonsum62
-TRANS=human40.cgi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NOM-ext -MILEN=0 -MAXLEN=200000000
-USER=US09508418.ecgn1_1_222 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPPX
-WAIT -THRAD=1

Search information block:

Query: US-09-508-418-2
Query length: 548
Database: N_Geneseq_0601.*
Database sequences: 730101
Database length: 313950809
Search time (sec): 112.980000

score_list:

| Sequence | Strd Orig | ZScore | EScore | Len | ! Documentation | .. | 2.0e-243 | 1874 |
|--|-----------|--------|---------|---------|-----------------|----|----------|------|
| /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAx28279 | | | 2836.00 | 4513.22 | | | 4.6e-243 | 1874 |
| /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAx28278 | | | 2832.00 | 4506.82 | | | 5.4e-179 | 1826 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04313 | | | 2113.00 | 3356.55 | | | 5.4e-179 | 1826 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04312 | | | 2113.00 | 3356.55 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04309 | | | 2077.00 | 3299.59 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04304 | | | 2077.00 | 3299.59 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04303 | | | 2077.00 | 3299.59 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04302 | | | 2077.00 | 3299.59 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04301 | | | 2077.00 | 3299.59 | | | 8.1e-176 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAV04300 | | | 2075.00 | 3296.57 | | | 5.1e-175 | 1719 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04299 | | | 2068.00 | 3285.19 | | | 1.5e-174 | 1632 |
| /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:AAV04298 | | | 2068.00 | 3276.95 | | | 1.8e-174 | 1847 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04309 | | | 2065.50 | 3275.61 | | | 1.8e-174 | 1847 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04308 | | | 2065.50 | 3275.61 | | | 1.8e-174 | 1847 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04307 | | | 2062.50 | 3273.61 | | | 1.6e-169 | 1784 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04306 | | | 2062.50 | 3273.61 | | | 1.6e-169 | 1784 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04305 | | | 2006.50 | 3186.37 | | | 1.6e-169 | 1784 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04304 | | | 2006.50 | 3186.37 | | | 7.3e-168 | 1911 |
| /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAV04303 | | | 1988.00 | 3155.01 | | | 9.2e-168 | 1811 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04302 | | | 1987.00 | 3155.01 | | | 9.2e-168 | 1811 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04301 | | | 1987.00 | 3155.01 | | | 9.2e-168 | 1811 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04300 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04299 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04298 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:AAV04297 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:AAV04296 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04295 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04294 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04293 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04292 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04291 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04290 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04289 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04288 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04287 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04286 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04285 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04284 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04283 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04282 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04281 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04280 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04279 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04278 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04277 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04276 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04275 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04274 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04273 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04272 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04271 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04270 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04269 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04268 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04267 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04266 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04265 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04264 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04263 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04262 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04261 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04260 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04259 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04258 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04257 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04256 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04255 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04254 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04253 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04252 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04251 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04250 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04249 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04248 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04247 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04246 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04245 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04244 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04243 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04242 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04241 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04240 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04239 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04238 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04237 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04236 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04235 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04234 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04233 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04232 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04231 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04230 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04229 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04228 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04227 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04226 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04225 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04224 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04223 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04222 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04221 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04220 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04219 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04218 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04217 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04216 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04215 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04214 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04213 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04212 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04211 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04210 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04209 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04208 | | | 1978.00 | 3141.35 | | | 5.3e-167 | 1691 |
| /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:AAV04207 | | | 1978.00 | | | | | |

|||||
176 TGCTCCGTTGCCAAAGATTACAGATTCTCTCAGCGGTCGACGGGG 225
67 yProAlaIaGIuLeuAspCysValIleValGlyAlaGlyIleSerGlyL 84
|||||
226 ACCGCCCGGAGCGTGGAGCTGTGTATATGTGAGAGCAAGAAATTGTGGCC 275
84 eucysIIaIaGIuValMetSerAlaAsnTyPProAsnLeuMetValThr 100
|||||
276 TGTGATTCGCCAGGTGATGTCCGCTAATTACCCCAATTGATGTAC 325
101 GIuAlaArgAspArgAlaGlyGlyAsnIleThrValGIuArgAspGI 117
|||||
326 GAGCGGAGAGATCGTCCGCTGGCAACATAAGACTGTGAAAGACAGCG 375
117 yTyLeuThrPGluGIuGIuGIuProAsnSerPheGlnProSerAspProMetL 134
|||||
376 CTATTGTGGGAAAGAGTCCCAACAGTTTCCACCGCGTCCGATCTATGT 425
134 eutHrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGIuAsp 150
|||||
426 TGAGTATGGCAGTAGATGTGTGATTTGAAGAGATTTGTGTGGAGAT 475
151 ProAsnAlaProArgPheValLeuThrLysGlyLysLeuArgProValPr 167
|||||
476 CCTAATGGCCCCGTTTGTGTGGAAAGGTAATTAAGCCCCGTCC 525
167 oSerLysLeuThrAspLeuProPhePheAspLeuMetSerLleProGlyL 184
|||||
526 CTCAAAATCTACGTACTTCCCTTTTGTGATTTATGACACTTCTGGCA 575
184 ySLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProGlyL 200
|||||
576 AGTTGAGAGCTGTGTTTGTGTCATTTGCCCTCCGCCCTTCACTCCAGGT 625
201 HtSGluGIuSerValGIuGlnPheValArgArgAsnLeuGlyGIuVal 217
|||||
626 CATGAGGATCAGTTGAGCAGTTGCGCTGTAATCTGTGTGCGAGAT 675
217 lPheGIuArgLeuIleGluProPheCysSerGIuValTyValGIuAspP 234
|||||
676 CTTTGACGCTTGATAGAACCATTTTGTCTGCGTTTATTTGTGTGATC 725
234 roSerLysLeuSerMetLysAlaIaIaPheGlyLysValITrplyLeuGIu 250
|||||
726 CCTCAAAACTGAGTATGAAGCAGCATTTGGAAAGTTTGAAGTTGGAA 775
251 GIuThrGIuGlySerIleIleGlyGlyThrPheLysAlaIleLysGIuAr 267
|||||
776 GAAACTGGTGTAGCATTTATTGGAAGAACCTTTAAAGCAATTAAGAGAG 825
267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
|||||
826 ATCCAGTACACCTTAAGCGCCCGGATCGCGTTTACCTTAACCAAAAG 875
284 lYgInThrValGIuSerPheArgLysGIuLeuArgMetLeuProAspAla 300
|||||
876 GACAGACAGTTGATCANTCAGGAAGGCTCAGAAATGCTCCGGAGTCA 925
301 lIleSerAlaArgLeuGIuSerLysLeuLysLeuSerTrplyLysLeuSer 317
|||||
926 ATCAGTGCMAAGATTGGCAACAAATTAACATATCATGCAAGCTTTCTAG 975
317 lIleThrLysSerGIuLysGIuGIuTyThrIleLeuThrTyGIuLthrProG 334
|||||
976 CATTACTAAGTCAGAAAAAGAGATATCATCTGACATAGCAGACACAG 1025
334 lUGlyValAlaSerLeuGlnSerArgSerIleValImetThrValProSer 350
|||||
1026 AAGAGTAGTTCTCTCAAGTGAAGCANTGTGATGACTGTGCCATCC 1075
351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAspAl 367
|||||

1076 TATGTAGCAAGCAACATATTACGTCTTTCGTTGCCGACAGATGC 1125
367 aluSerAsnPheTyTyPProProValGIuAlaValIThrIleSerTyP 384
|||||
1126 ACTTCAAAATTTCTACTATGCCAGTGTGAGCAGTCAACAATTTCTATTC 1175
384 roGIuGIuAlaIleArgAspGIuArgLeuValAspGIuLeuLysGly 400
|||||
1176 CTCAGAGGTATTCGTGATGAGCGTCTGTGATGTGATCACTAAAGGA 1225
401 PheGIuGIuLeuHtSProArgThrGIuGIuValGIuThrLeuGIuThrII 417
|||||
1226 TTTGGCGAGTTGTCATCCAGTACACAGGAGTGGAAACCTAGGAACGAT 1275
417 eTySerSerSerLeuPheProAsnArgAlaProLysGIuArgValLeuL 434
|||||
1276 ATATAGTTCATCCTCTTCCCTAACCGTCCCAAAAGGTGGGTGCTAC 1325
434 euleuAsnTyRIleGIuAlaLysAsnProGIuITrleuSerLysThr 450
|||||
1326 TCTTGAACCTACATTGGAGGAGCAAAAATCCGAAATTTTGTCTAAGACG 1375
451 GIuSerGIuLeuValGIuValValAspArgAspLeuArgLysMetLeuII 467
|||||
1376 GAGAGCCAACTGTGGAGATGATGTGATCGTACCTCAGAAAAATGCTTAT 1425
467 eLysProLysAlaGIuAspProLeuValGIuValArgValITrProG 484
|||||
1426 AAAACCCAAAGCTCAAGATCCTTGTGTGGTGTGGCAGATGTGCCAC 1475
484 lAlaIleProGIuPheLeuValGIuHisLeuAspThrLeuSerThrAla 500
|||||
1476 AACGTATCCACAGTTTTGTGTGTCATCTGATCGATACCTAAGTACTGCA 1525
501 LysAlaAlaMetAsnAspAsnGIuLeuGIuGIuPheLeuGIuGlyLys 517
|||||
1526 AAAGCTGTATGATGATATATAGGCTTGAAGGCGTGTCTTGGGGGGTAA 1575
517 nTyValSerGIuValAlaLeuGIuArgCysValGIuGIuAlaTyGluV 534
|||||
1576 TTTATGTCTCAGGTGATGCATTTGGGAGGTGTGTGAAGTGTCTTATGAAG 1625
534 alAlaSerGIuValIThrGIuPheLeuSerArgTyAlaTyLys 548
|||||
1626 TTGCATCCGAGTAAACAGATTTCTGTCTCGGATGCATACAA 1669
seq_name: /SID58/gcdata/geneseq/geneseq/NA1999.DAT:AAx28278
seq_documentation_block:
ID AAx28278 standard; cDNA to mRNA; 1874 BP.
XX
AC AAx28278;
DT 16-JUN-1999 (first entry)
XX
DE N. tabacum protoporphyrinogen oxidase coding sequence.
XX
KW protoporphyrinogen oxidase; light-regulating herbicide tolerance; enzyme;
KM ds.
XX
OS Nicotiana tabacum.
XX
PN W09913087-A1.
XX
PD 18-MAR-1999.
XX
PF 10-SEP-1998; 98WO-JP04064.
XX
PR 11-SEP-1997; 97JP-0265084.
XX
PA (NIPPON NOYAKU CO LTD.
XX
Hirooka T, Horikoshi M, Mametsuka K;
PI

XX WPI: 1999-215067/18.
 DR P-PSDB: AAY05200.
 XX
 PT Protoporphyrinogen oxidase tolerant to light-requiring herbicides,
 useful for constructing plants with such properties
 XX
 PS Example 5: Page 34-39; 56pp: Japanese.
 XX

This sequence encodes the Nicotiana tabacum protoporphyrinogen oxidase
 of the invention, which is tolerant to light-requiring herbicides. The
 enzyme and its derivatives can be applied to construct plants with high
 CC tolerance to light-requiring herbicides, useful in agriculture.
 CC
 XX

Sequence 1874 BP: 509 A; 383 C; 461 G; 521 T; 0 other;

alignment_scores:
 Quality: 2832.00 Length: 548
 Ratio: 5.177 Gaps: 0
 Percent Similarity: 99.818 Percent Identity: 99.635

alignment_block:
 US-09-508-418-2 x AAY28278 ..

Align seg 1/1 to: AAX28278 from: 1 to: 1874

```

1 MetThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnse 17
  |||||||
26 ATGACACAACTCCCATGCCAATCATCTTAATATTTCTACCTCACAGTC 75
  |||||||
17 rSerSerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPhes 34
  |||||||
76 GTGCTATCGCCCATTTGGCATTTCTTAACCGTAGAGTTTCATCCCTTTC 125
  |||||||
34 eSerIleSerIysArgAsnSerValAsnGlySasnGlyTrpArgThrArg 50
  |||||||
126 CTTCACATCTCCAAAGCGCAATGATGTCATATTGCCAATGGCTGGAGAACACA 175
  |||||||
51 CysSerValAlaIysAspTyrThrValProSerSerAlaValAspGly 67
  |||||||
176 TGCTCGCTTCCCAAGATTACACAGTTCCCTCTCAGCGGTCCACGGCGG 225
  |||||||
67 yProAlaAlaGluLeuAspCysValIleValAlaGlyAlaGlyIleSerGly 84
  |||||||
226 ACCGCGCGCGAGCTGGACTGTGTATAGTTGAGCAGGAATTAAGTGGCC 275
  |||||||
84 eucysIleAlaGlnValMetSerAlaAsnTyrProAsnLeuMetValThr 100
  |||||||
276 TCTGCATTTGGCAGGTGATGTCGCTAATTATCCCAATTTGATGTAAACC 325
  |||||||
101 GluAlaArgAspArgAlaGlyIleAsnIleThrThrValGluArgAspG 117
  |||||||
326 GAGCGGAGATATCGTCCGCGTGGCAACATACACATCTGTGAAAGACAGCG 375
  |||||||
117 yTyrLeuTrpGluGlnGlyProAsnSerPheGlnProSerAspProMet 134
  |||||||
376 CTATTGTGGGAGAGAGTCCCAACAGTTTCCAGCGTCCGATCTATGT 425
  |||||||
134 eutThrMetAlaValAspCysGlyLeuIleAsnAspAspLeuValIleuGlyAsp 150
  |||||||
426 TGACTATAGGAGATGATTTGGATTTGAAGATGATTTGGTGGAGAT 475
  |||||||
151 ProAsnAlaProArgPheValIleuTrpLysGlyLysLeuArgProValPr 167
  |||||||
476 CCGAATCGCGCCGCTTTCGTTTGTGGAAGGTAATAATTAAAGCCCGCTCC 525
  |||||||
167 oSerLysLeuThrAspLeuProPheAspLeuMetSerIleProGlyL 184
  |||||||
526 CTCAAACACTACTGATCTCTCTTTTGTGATTTGATGAGCATTCCTGGCA 575
  |||||||
184 yIleuArgAlaGlyPheGlyProIleGlyLeuArgProSerProProGly 200
  |||||||

```

```

576 AGTTGAGAGCTGTTTTGGTCCATTTGGCTCCGCCCTTCACTCCAGGT 625
  |||||||
201 HIsGluIleSerValGluGlnPheValArgArgAsnLeuGlyGlyVal 217
  |||||||
626 CATGAGGAATCAGTTGACAGATGTCGTCGTATCTTGGTGGCGAAGT 675
  |||||||
217 lPheGluArgLeuIleGluProPheCysSerGlyValTyrValGlyAsp 234
  |||||||
676 CTTCAGCGCTGATAGAACCATTTTGTGCTGTTATCTGCTGATGTC 725
  |||||||
234 roSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLeu 250
  |||||||
726 CTCAAACCTGAGTATGAAGCAGCATTTGGGAAAGTTTGGAGCTTGGAA 775
  |||||||
251 GluThrGlyGlySerIleIleGlyLysThrPheLysAlaIleLysGluArg 267
  |||||||
776 GAAACTGGTGTAGCATTTTGGAGAACCTTTAAGCAATTAAGGAGAG 825
  |||||||
267 gSerSerThrProLysAlaProArgAspProArgLeuProLysProLysG 284
  |||||||
826 ATCCAGTACACCTTAAGCGCCCGGATCCGGTTTACCTTAACCAAAAG 875
  |||||||
284 yGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAspAla 300
  |||||||
876 GACAGACAGTTGGATCATTCAGAGAGGCTCAGAAATGCTCCGATGCA 925
  |||||||
301 lIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLysLeuSer 317
  |||||||
926 ATCAGTGCAGATTTGGAAAGCAAAATTAAACTATCATGGAACCTTTAG 975
  |||||||
317 rIleThrLysSerGlyLysGlyLysGlyLysLeuThrTrpGluThrProG 334
  |||||||
976 CATTACTAAGTCAGAAAAAGAGGATATCACTTGACATACAGACAGACAG 1025
  |||||||
334 lGluValValSerLeuGlnSerArgSerIleValMetThrValProSer 350
  |||||||
1026 AAGGAGTGTCTCTTCAAACTCGACATTTGTATCTGCTGCTGCTATCC 1075
  |||||||
351 TyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAlaAspAl 367
  |||||||
1076 TATGTAGCAGACACATATTACGTCTCTTTCGGTTGGCAGCAGAGATC 1125
  |||||||
367 aLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSerTyrP 384
  |||||||
1126 ACTTTCAAAATTTCTACTATCCCGCCAGTTGGAGCAGCTCAAAATTCATATC 1175
  |||||||
384 roGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLysGly 400
  |||||||
1176 CTCAGAAAGCTATTTGCTGATGAGCGTCTGTTGATGTGAACCTAAGGGA 1225
  |||||||
401 PheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyThrI 417
  |||||||
1226 TTGGGCAAGTTGCATCCACGTACACAGGAGTGAACACTAGGAACAT 1275
  |||||||
417 eTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgValIleu 434
  |||||||
1276 ATATATGTCATCACTCTTCCCTTAACCGTGGCCCAAAAGTGGGGCTAC 1325
  |||||||
434 euleuAsnTyrIleGlyGlyAlaLysAsnProGluIleLeuSerLysThr 450
  |||||||
1326 TCTTGAACTACATTGGAGAGCAAAAATCTCGAATTTTGTCTAAGCG 1375
  |||||||
451 GluSerGlnLeuValGluValAlaValAspArgAspLeuArgLysMetLeu 467
  |||||||
1376 GAGAGCAACTGTGGAGAGTACTGATGCTGACCTCAGAAAAATCTTAT 1425
  |||||||
467 eLysProLysAlaGlnAspProLeuValValGlyValArgValTrpProG 484
  |||||||
1426 AAAACCCAAAGCTCAAGATCCCTTGTGTGGGTGGAGATGGCCAC 1475
  |||||||
484 lAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerThrAla 500
  |||||||
1476 AAGCTATCCACAGTTTGTGGTGCATCTGATACCTAAGTACTGCA 1525
  |||||||

```

| | | |
|-------------------------------|--------------------------|--|
| | seq_name: | /SID58/gcgcdata/geneseq/geneseqn/NA1997.DAT.AAV04313 |
| | seq_documentation_block: | |
| XX | ID | AAV04313 standard; cDNA; 1826 BP. |
| AC | NC | AAV04313; |
| XX | DT | 20-Apr-1998 (first entry) |
| XX | DE | Cotton protox-1 cDNA. |
| XX | OS | Gossypium hirsutum. |
| XX | FH | Key |
| FT | CD5 | Location/Qualifiers 31..1650 /*lag= a /product= protox-1 |
| XX | PN | M09732028-A1. |
| XX | PD | 04-SEP-1997. |
| XX | PE | 27-FEB-1997; 97MO-US03343. |
| XX | PR | 21-JUN-1996; 96US-0020003. 28-FEB-1996; 96US-0012705. 28-FEB-1996; 96US-0013612. |
| XX | PA | (NOVS) NOVARTIS AG. |
| XX | P1 | Johnson MA, Volrath SL, Ward ER; MPI; 1997-489209/45. P-PSDB: AAMW1609. |
| XX | DR | DNA containing a plant proto-porphyrinogen oxidase gene promoter - optionally linked to a heterologous gene, especially to express herbicide-resistant enzymes, and plants containing such constructs |
| XX | PS | Disclosure; Pages 78-80; 114pp; English. |
| XX | CC | The present sequence encodes cotton protoporphyrinogen oxidase-1 (protoc-1). |
| CC | CC | The protoc-1 promoter can be used to express herbicide resistant enzymes, specifically protoc, i.e. a plant tissue, plant or progeny containing a chimeric gene of the promoter and a heterologous coding sequence. The plant can also be used in breeding programmes. Also hybridising fragments of the protoc coding sequence can be used as probes, e.g. to isolate related genes or for genomic mapping. |
| SQ | Sequence | 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other; |
| alignment_scores: | | |
| Quality: 2113..00 Length: 506 | | |

| | | | | | | | |
|---|--|-------------------|--|---|-------|--|--|
| Percent Similarity: | 99.281 | Percent Identity: | 79.249 | Ratio: | 4.477 | Gaps: | 1 |
| Alignment block: | | | | | | | |
| US-09-508-418-2 x AAV04313 | | | | .. | | | |
| Align seg 1/1 to: AAV04313 from: 1 to: 1826 | | | | | | | |
| 47 | TRPAQTlThrArgSerValAlaLysAspYrThrValProSerSerAl | 63 | avalAspIglYlProAlaAlaIglLeuAspCysValIlleValGlYlaG | 80 | 180 | AATCGACGGGGAGAAATCATCATCCGCGATGGGTGATCTGTGGAGGTG | 222 |
| 130 | TTCAAGCTCCGATGTCTCCCTGGCCGAGGGTCCCAAGATTCTCATCTAA | 179 | TTTGGAAAGATGTGAAGCTTAAGAAATGTGTGGCAGCATCTGGTGG | 719 | 730 | TTTGGAAAGATGTGAAGCTTAAGAAATGTGTGGCAGCATCTGGTGG | 779 |
| 243 | PhcglYlYSValITpIlylSleuGlulThcglYlSerIlleIleglYl | 259 | YThrPhcYlSAlAllelYsGlYlgrSerSerThProLYsAlAProArg | 276 | 780 | CACITTCAGACATCCAGAGAGAAATTAAGACCTTAACCCACCCAGAG | 829 |
| 276 | sPrOAlgrLeuProlYsProlYsIglInThrValIglYSerPheArqLys | 322 | sPrOAlgrLeuProlYsProlYsIglInThrValIglYSerPheArqLys | 322 | 293 | GLYLeuArqMeLleuProlAspAlIlleSerAlArqLeuIglYSerLysIle | 309 |
| 830 | AACCGCGCTCTCCAAACCGAAGGCGCAACAGTTGATCTTTTAGAGAG | 879 | 880 | GGACTTACCATGCTGCTGAGGCAATTGCTTAACAGTTGGGTAGCAATGT | 9299 | 309 | uLysLeuSerTrpLysLeuSerSerIlleThrLysSerGluLysGlYlGT |
| 309 | uLysLeuSerTrpLysLeuSerSerIlleThrLysSerGluLysGlYlGT | 326 | | | | | |


```

|||||
280 GTCGGTTCATGATGATGTCACGAGAGCCGAGACCGTGTGGTGCAA 329
109 nIleThrThValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn 126
330 CATCACTACCGCTGAGAGATGATATCTGTGGAGAAAGGCCCAACA 379
126 erpHeGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
380 GTTTCACCCCTCCGATCTTCTAACCATGAGCCGTGATGATGATG 429
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeu 159
430 AAGACGATTTGGTTTAGTGCACCCCTAATGCACCGCATTTGACATG 479
159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
480 GGAGGAAACTAAAGCCCTGTGCCCTCCAAAGCCAAAGCTTGCCTTT 529
176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
530 TTGATTGATGACATTCGTGAAAACCTTAAGGCTGGGCTGGGGCTATT 579
193 GlyLeuArgProSerProGlyHisGluGluSerValGluGlnPheVal 209
580 GGCAATGCGCCCTCCCGCTCGGGGTATGAAAGATCGGTGAGAGATTGT 629
209 lArgArgAsnLeuGlyGlyValPheGluArgLeuIleGluProPheC 226
630 GCCCGCTAATCTGTGCTGAGGTTTGAACGCTTATTTGAAACCAATTT 679
226 ySserGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
680 GTTCAGGGTTTATGACAGGGATCTTCAAAATTAAGCATGAAGCCAGCA 729
243 PheGlyLysValTyrPylsLeuGluGluThrGlyLysSerIleIleGly 259
730 TTTGGAAGATGATGAGACGTAGAAAGATGTGTGCACATCATTTGGTGG 779
259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
780 CACTTTCAGACAAATCCAGAGAGAAATTAAGACCTTAAGCCACCAGAG 829
276 sPProArgLeuProLysProLysGlyInThrValGlySerPheArgLys 292
830 ACCCGCTCTGCCAAACCGAAGGCCAAACAGTTGGATCTTTAGAGAG 879
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
880 GGACTTACCATGCTGCTGAGGCAATTCCTAACAGTTTGGGTACCAATGT 929
309 uLysLeuSerTyrPylsLeuSerSerIleThrLysSerGluLysGlyT 326
930 AAAATTATCTTGGAGCTTTCAGATTACCAAAATGGCGAATGAGAGGT 979
326 yThsLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
980 ATAACTTGACATTTGAACACCTGAGAGAAATGATATCTTCAAGATGA 1029
343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
1030 AGTGTGTAATGACATTCATCCATGCTGTCCAGTAACCTTGTCATCC 1079
359 cLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProPro 376
1080 TCTCTCGGCTCTGCTGCAGATGATTAATCCAAATTTATATATCTCCAG 1129
376 aGlyAlaValThrIleSerTyrProGluGlnAlaIleArgAspGluArg 392
1130 TTGCATCGTACACAGTCTCTATCCAAAGAGACCATTCGAAAGAAATGT 1179
393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrG 409
|||||

```

```

1180 TTGATTGATGTGTAACCTTAAGGGGTTTGGCCAGTTGCACCCAGCAGCCA 1229
409 nGlyValGluThrLeuGlyThrIleTyrSerSerLeuPheProAsn 426
1230 AGGAATTGAACCTTAGGAGCATATACAGTTCACTCATCTTCCCAATC 1279
426 rGluProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
1280 GAGCTCATCTGCGACAGGTGCTCTTGAACACTACATAGAGAGACTACC 1329
443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValAs 459
1330 AACATGCAATTTTGTCCAAAGCTGAAGGGCAACTGTGTAAGACAGTTGA 1379
459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476
1380 TCGTGATTTGAGAAAAATGCTTAATATCTTAATGCAAAAGATCCCTTG 1429
476 aLValGlyValArgValTyrProGlnAlaIleProGlnPheLeuValGly 492
1430 TTTTGGGTGTAAAGATATGCCCCAAAGCCATTCACAGTTCTTGGTGT 1479
493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
1480 CATTTGATCTCCTGTGATGATGCAAAATGCGCTCAGGAGATTCTGGGTT 1529
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGly 526
1530 TCATGGACGTTTCTTGGGGCAACTATATCTGTGTGGCATTAAGGAC 1579
526 rGcysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
1580 GGTGTGTGGAAGTGTCTACGAGGTTGACAGCTGAAGTGAAGAAATCTCTG 1629
543 SerArgTyrAlaTyrLys 548
1630 TCACAATATGCAATACAA 1647

seq_name: /SID8/gcdata/geneseq/geneseq/NA2001.DAT:AAF76579
seq_documentation_block:
ID AAF76579 standard; cDNA: 1826 BP.
XX
AC AAF76579;
XX
DT 16-MAY-2001 (first entry)
XX
DE Cotton protoporphyrinogen oxidase coding sequence SEQ ID NO: 15.
XX
KW Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;
KW soybean; sugar beet; oilseed rape; sugar cane; mutant; mutein; ss.
XX
OS Gossypium hirsutum.
XX
PN WO200112825-A1.
XX
PD 22-FEB-2001.
XX
PF 30-JUN-2000; 2000WO-EP06127.
XX
PR 13-AUG-1999; 99US-0373691.
XX
PA (SYNG-) SYNGENTA PARTICIPATIONS AG.
XX
PI Johnson MA, Volrath SL, Helfetz PB, Law MD;
XX
DR WPI: 2001-234914/24.
XX
P-PSDB: AAB72907.
XX
PT Plant DNA molecules encoding herbicide-tolerant forms of
PT protoporphyrinogen oxidase which are useful for rationally designing
PT new inhibitory herbicides and for producing herbicide-tolerant
PT transgenic plants and seeds -

```

```
XX
PS Claim 2; Page 180-184; 228pp; English.
CC
XX The present invention provides the protein and coding sequences of a
CC number of herbicide-tolerant forms of wheat, soybean, rice, sorghum,
CC sugar beet, sugar cane, cotton and oilseed rape protoporphyrinogen
CC oxidase (protox) enzyme. Examples of these mutants are shown in
CC AAB72920-AAB72926. They are useful as they enable the production of
CC herbicide-tolerant plants and seeds. The present sequence is a protox
CC coding sequence.
XX
SO Sequence 1826 BP; 471 A; 399 C; 439 G; 517 T; 0 other;

alignment_scores:
    quality: 2113.00      Length: 506
    Ratio: 4.477          Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:
US-09-508-418-2 x AAF76579 ..
Align seg 1/1 to: AAF76579 from: 1 to: 1826

47 TPrArGThrArgCysSerValAlaLysAspLysThrValProSerSerAl 63
   ::::: |||||:::|||||:::|||||:::|||||:::|||||
130 TTCAAGCTCCGATGCTCTCCCTGCGGAGGTCGCCAGATTCCTCTAA 179
   ::::: |||||:::|||||:::|||||:::|||||:::|||||
63 aValAspLysGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
   ::::: |||||:::|||||:::|||||:::|||||:::|||||
180 AATGACGAGGGGAGATCATCATCGGAGTTGGTCATCGTTGGAGGTG 229
   ::::: |||||:::|||||:::|||||:::|||||:::|||||
80 LylIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
230 GTATCAGTGGCCTTGATGCTCAACGCTCGCGCACCAAGCACCGTGAC 279
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
280 GTCCCTCCCATGTGATGTGACGAGGCCAGACCGTGTGGTGCAA 329
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
109 nIleThrThrValGluArgAspGlyTyrLeuTyrGluGluGlyProAsn 126
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
330 CATCAGTACCGTTGAGAGATGATATCTGTGGAGAAAGGCCCCACA 379
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
380 GTTTTCAGCCCTCGATCTTCTTAACCATGCGCGTATGATGATG 429
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
430 AAGGACGATTTGGTTTGGTGAACCTTAATGACCCGGATTTGTACTATG 479
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
480 GGAGGGAATACTAAGGCGTGTGCCCTCCAGCAACCGACTTCCGCTTT 529
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
530 TTGATTTGATGACCATTTGCGAAACTTGAAGGCTGGGCTCGGGCTATT 579
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
193 GlYleuArgProSerProProGlyHisGluSerValGluGlnPheVal 209
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
580 GGCATTCGGCTCCCTCCGCGGTTTGAAGAAATGCGTGGAGGATTTGT 629
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
209 LArgArgAsnLeuGlyGlyValPheGluArgLeuIleGluProPheC 226
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
630 GCGCCGTAATCTGTGCTGAGGTTTGAACGCTTATTGAACCATTTT 679
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
226 ySerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIle 242
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
680 GTTCAGGTGTTTATGACGGGATCTTCAAAATTAAGCATGAAGACGACA 729
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
243 PheGlyLysValTyrPlyLeuGluGlnThrGlySerIleIleGlyGly 259
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||

730 TTGGAAGAGTATGAGAACTAGAAAGATTGGTGGCAGCATCATTTGGT 779
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
259 yThrPheLysAlaIleLysGluArgSerThrProLysAlaProArgA 276
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
780 CACTTTCAGAACATCCAGAGAAATTAAGACCTTAAGCCACCGAGAG 829
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
276 sProArgLeuProLysProLysGlyIleThrValGlySerPheArgLys 292
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
830 ACCCGGCTCGCCAAAACCGAAGGCCAAACAGTTGATCTTTAGAAAG 879
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
880 GGACTTACCATGCTCGCTGAGCAATGCTAAGACGTTGGTAGCAATGT 929
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
309 uLysLeuSerThrPlyLeuSerSerIleThrLysSerGlyGlyTyr 326
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
930 AAATTAATCTTGGAAAGCTTCCAGTATTAACAAATTGGCAATGAGAGGT 979
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
326 yHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
980 ATACCTTGACATTTGAACACCGTGAAGATGTATCTCTCAGAGTACA 1029
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1030 AGTGTGTATGTACCATTCATCCATGTCAGTACAGTATGTTGATC 1079
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProPoy 376
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1080 TCTCTCGCTGCTGCTGCGAGATGATATCCCAATTTATTTCTCCAG 1129
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
376 aLeGlyAlaValThrIleSerTyrProGluGluAlaIleArgAspLuar 392
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1130 TTGATGATGTGTGAACCTTAAGGGGTTGGCCAGTTGCACCCAGCA 1179
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
393 LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1180 TTGATGATGTGTGAACCTTAAGGGGTTGGCCAGTTGCACCCAGCA 1229
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1230 AGGAATGGAACCTTATAGGACATATACGTTATCATCTTTCCCAATC 1279
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
426 rGAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyValAlaLys 442
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1280 GAGCTCCATCTGCGAGGGGTGCTCTTGAATCATATGAGAGACTACC 1329
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
443 AsnProGluIleLeuSerLysThrGluSerGlnLeuValGluValAlas 459
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1330 AACACTGGAATTTTGTCCAAAGACTGAAGGGGAACCTGTAGAAGCAGTGA 1379
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1380 TCGTATTTGACAAAATATCTTATTAATCTTAATGCAAGATCTCTTGG 1429
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
476 aValGlyValArgValTyrProGluAlaIleProGlnPheLeuValGly 492
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1430 TTTTGGGTGTAGAGATAGGCCAAAAGCCATTCACAGTTGTGGTGGT 1479
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
493 HisLeuSerThrLeuSerThrAlaLysAlaIleMetAsnAspAsnGlyLe 509
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1480 CATTTGGATCTCCCTTGATAGTGCAAAATATGCTCTCAGAGGATTCGGGTT 1529
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
509 uGluGlyLeuPheLeuGlyLysAsnTyrValSerGlyValAlaLeuGlyA 526
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1530 TCATGAGACTGTTCTTGGGGGCAACTATGTAATCTGATGATTAAGGAC 1579
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
526 rGysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 542
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1580 GGTGTGTGAAGGTCTTAAGAGGTTGCAAGCTGAAGTAAGAAATTCCTG 1629
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
543 SerArgTyrAlaTyrLys 548
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

1630 TCACATATGTCATACAA 1647

seq_name: /SID8/gcgdata/geneseq/geneseqn/NA1996.DAT:AA11673

seq_documentation_block:

ID AA11673 standard; cdna; 1719 BP.

AA11673:

02-JUL-1996 (first entry)

DNA encoding a protein having protoporphyrinogen oxidase activity.

Protoporphyrinogen oxidase; PPO; herbicide; resistance; mutant

variegate porphyria; ss.

Arabidopsis sp.

Key Location/Qualifiers

FT CDS 31, 1644

/product= Protoporphyrinogen oxidase.

MO9534659-A1.

21-DEC-1995.

08-JUN-1995; 95WO-IB00452.

16-JUN-1994; 94US-0261198.

(CIBA) CIBA GEIGY AG.

Volrath S, Ward ER;

WPI; 1996-049687/05.

P-PSDB; AAR90295.

DNA encoding eukaryotic proto:protoporphyrinogen oxidase and herbicide

resistant mutants - used to make herbicide resistant plants and for

diagnosis and treatment of variegate porphyria

Claim 1; Page 69-71; 118pp; English.

DNA encoding a modified protoporphyrinogen oxidase (PPO) or a

chimeric gene comprising a promoter (pref. active in a plant) linked

to such a DNA or DNA encoding a wild type PPO can be used (1) to

impart herbicide resistance to plants; (2) for treating and

diagnosing deficient PPO activity in animals (esp. variegate

porphyria); and (3) for the production of recombinant PPO which is

useful as an assay reagent and in rational design of new inhibitory

herbicides. Herbicide resistant PPO genes can also be used to

select plants transformed with a transgene and probes derived from

the genes can be used to quantify levels of PPO mRNA.

Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;

alignment_scores:

Quality: 2077.00

Ratio: 4.230

Percent Similarity: 91.095

Length: 539

Gaps: 4

Percent Identity: 75.139

alignment_block:

US-09-508-418-2 x AA11673 ..

Align seg 1/1 to: AA11673 from: 1 to: 1719

```

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
28 TCACATGAGATTATCTCTCTCGTCGACGAGACTCAATGCTTCTCCGTC 77
35 rHisLeuThrTyrgluThrProgluGluValSerLeuGlnSerArgS 49

```

```

|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTATAGCCTTATAGC 124
49 hrArgCysSerValAlaValAspTyrThrValProSerSerAlaValAsp 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
125 TCCGTTGTTTACAGTCGCGCGGTGACCAACCCGCGATCTTCAAAATCGAA 174
66 GLYGLYProAlaIleValLeu.....AspCysValIleValAlaIle 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
175 GCGCGAGAGGCGCACCATCACAGAGGATTGTGTATTTCTGGCGCGAG 224
80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
225 TATTAGTGGCTTGTGATCGCTCAGCGCGCTGCTACTAAGCATCTCGAG 274
95 ....ProAsnLeuMetValThrGluAlaArgAspArgIleValGlyAsn 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
275 CTGCTCCGAAATTAAATGTGACCGAGGCTAAGGATGCTTGGAGGCAAC 324
325 ATTATCAGCTCGTAGAGAAATGTTTCTCTGGAAGAAAGTCCCATAG 374
126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
375 TTTTCAACCGTCTGATCTATGCTCATATGCTAGTGTAGATAGCTTTGA 424
143 ysAspAspLeuValIleuGlyAspProAsnAlaProArgPheValLeuTrp 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
425 AGGATGATTTGGTGGGAGATCTACTGCGCAAGTTGTGTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheh 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
475 AATGGGAATTTGAGCGCGGTTCATCGAAGCTTAACAGACTTACCGTCTT 524
176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
525 TGATTGATGATGATTTGGTGGAGATTAAGCTGTGTTGGTGCACCTTG 574
193 LysLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVal 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
575 GCATTCAGCCGTCACCTCCAGCTCGTAGAAGAAATCTGGAGAGTTTGA 624
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
625 CGCGGTAACTCGGTGATGATGATTTTGAAGCGCTTATGAAACCGTTTG 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
675 TTCAGGTTTATGCTGTGATCTTCAAAACGTGACATGAAAGCAGCGT 724
243 heGlyLysValTyrPheLysLeuGluGlnThrGlySerIleIleGlyGly 259
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
725 TTGGGAAGTTTGGAAACTAGACCAAAATGCTGGAACATATAGGTGGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgS 276
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
775 ACTTTAAGGCAATTCAGAGAGGAAACGCTCCCAAGGCAACAAGAGA 824
276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLysG 293
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
825 CCGCGGCTGCGCAAAACACAGGCGCCAAACAGTTGGTCTTTAGAGAAAG 874
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
875 GACTTGAATGTTGCCAGAGCAATATCTGCAAGATTAGTACCAAGTT 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGluLysGlyGly 326
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
925 AAGTTGTTGGAAGCTCTCAGATCATCACTGAGCTGAGCGGAGATA 974
326 rHisLeuThrTyrgluThrProgluGluValSerLeuGlnSerArgS 343
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


herbicide resistant mutants, useful to prepare plants resistant to herbicide which therefore kills undesired vegetation only

Example 2; Page 95-99; 196pp; English.

This cDNA clone codes for Arabidopsis protoporphyrinogen oxidase (Protox-1, see AAM25746). The clone has been deposited in pMDC-2 as NRRL B-21238. The isolated cDNA can be mutated so that it encodes a modified protox enzyme that is resistant to protox inhibitors and hence is herbicide tolerant. Plants, especially crop plants, may be engineered for resistance to protox inhibitors via mutation of the native protox gene to a resistant form, or they may be transformed with a gene encoding an inhibitor-resistant form of a plant protox enzyme, including claimed forms from wheat, soybean, cotton, sugarcane, oilseed rape, rice, sorghum and Arabidopsis (see AAM25738-48). Application of herbicide therefore kills undesired vegetation only. Plant protox nucleic acids can also be used as probes and PCR primers, as selectable markers in plant cell transformation methods, and for recombinant production of protox enzymes in host cells. Arabidopsis Protox-1 cDNA has been used in the identification of soybean (see AAT86122), sugarbeet (see AAT86124) and rape (see AAT86125) Protox-1 cDNA.

Sequence 1719 BP; 460 A; 351 C; 433 G; 475 T; 0 other;

alignment_scores:

Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

alignment_block:

US-09-508-418-2 x AAT86129 ..

Align seg 1/1 to: AAT86129 from: 1 to: 1719

```

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
28 TCCATGAGATTATCTCTCCGTCGACGACTCAATCGCTTCTCCGTC 77
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
35 rIleSerLysArgAsnSerValAsnCysasnGlyTrp.....ArgT 49
   |:::||||| ||| :|||: ||| :|||: ||| :|||: ||| :|||:
78 GTTTTCGAAGCCCAAT...CTCCGATTAAATGTTTAAAGCCTCTTAGAC 124
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
49 hArgCysSerValAlaLysAspTyrThrValProSerSerValAlaLys 65
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
125 TCCGTTGTCAGTGGCGGCTGGACCAACGCTGCATCTCAAAAATCGAA 174
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
66 GlyGlyProAlaIaIaGluLeu.....AspCysValIleValGlyAla 80
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
175 GCGCGAGGAGGACCAACCATCAGCAGGATGTGTGATTGTGCGGAGAG 224
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
80 yIleSerGlyLeuGlyIleAlaGluValMetSerAlaAsnTyr..... 94
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
225 TATTAGTGGTCTTGTGATCGTCAGCGCTGTGCTACTAGCAATCCTGATG 274
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
275 CTGCTCCGCAATTAATTGTGACCGAGGCTAAGATCGTGTGAGGCAAC 324
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
110 IleThrThrValGluArgAspGlyTyrIleuTrpGluGluGlyProAsn 126
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
325 ATTATTCACCTCGTGAAGAGATGTCTTCTGTGGAAAGAGTCCCAATAG 374
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 143
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
375 TTTTCAACCGCTGTGATCTCTATCTCACTATGCTGTAGTACTGGTTGA 424
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
143 yAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
425 AGGATGATTTGGTGGAGATCTACTAGCCCAAGTTTGTGTGG 474
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176

```

```

475 AATGGAAATTGAGCCGGTCCATCGAACCTAACACTTACCGTCTT 524
   ::|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
176 eAspLeuMetSerLleProGlyLysLeuArgAlaGlyPheGlyProIle 193
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
525 TGATTTGATGATGATTGTTGGAGAGATTAGACTGCTTTGGTGCACCTG 574
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
193 LysLeuArgProSerProProGlyHisGluLeuSerValGluGlnPheVal 209
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
575 GCATTCGACCGTCACTCCAGGCTGAGAGAAATCTGTGGAGAGTTGTA 624
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleLeuProPhe 226
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
625 CCGGGAATCCCGGTGTGAGGTTTGGACGCGCTGATTAACCGTTTG 674
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaLap 243
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
675 TTCAGGTGTTTATCTGCTGATCTTAAACTGACATTAAGACAGCCT 724
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
243 heGlyLysValTyrPlysLeuGluGluThrGlySerLleIleGlyGly 259
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
725 TTGCGAAGTTTGGAACCTAGACCAAAATGTGGAAACATATAGTGCT 774
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
775 ACTTTAAGCAATTCAGAGAGAGAAAAACGCTCCCAAGCAGACAGA 824
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
276 pProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 293
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
825 CCGCGCGCTGCCAANACACAGGCCAANACGTTGGTCTTTCAGAGAG 874
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLys 309
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
875 GACTTCGAATGTTGCCAAGCAATATCTCAAGATTAGGTAGCAAAAGT 924
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
310 LysLeuSerTrpLysSerSerLleThrLysSerLysGlyGlyTyr 326
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
925 AAGTTGCTTGGAACTCTCAGTATCATAAGCTGAGAGCGGAGATA 974
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
326 rHisLeuThrTyrGluThrProGluGlyValIleSerLeuGlnSerArg 343
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
975 CAATTAACATATGACATCCAGATGTTAGTTCGCTGCAGAGCAAA 1024
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1025 GTTGTGAATATGACGCTGCATCTCATGTGCAAGTGTCTTGCAGCCT 1074
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
360 LeuSerValAlaAlaIaAspAlaLeuSerAsnPheTyrTrpProPro 376
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1075 CTTCCTGAATCTGCTGCANATGACCTCTCAAAACTATATTTACCCAC 1124
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 393
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1125 TGCAGCAGATATCTATCTGTRACCCGAAAGAAAGCAATCCGAAAGAT 1174
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1175 TGATAGATGAGTGAACCTAAGGGTTTGGCAATTCATCCAGCAGACCA 1224
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsn 426
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1225 GGAATGTAACATTAAGAACTATCTACAGCTCCCTCATCTTCCAAATCG 1274
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyValAlaLys 443
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1275 CGCACCCCGCGAAGAAATTTTGTGTTGAACCTACATGCGGCGCTCA 1324
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
443 snProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAla 459
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1325 ACACCGCAATTCCTCCAGTCTGAAGGTGATGAGGAAGAGATTGAC 1374
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```


226 ssergIValTyrValGlyAspProSerLysSerMetLysAlaIap 243
|||||
675 TTCAGGTGTTTATAGCTGGATGCTTCAAACTGAGCATGAAAGAGCGCT 724
243 heGlyLysValTrrPlyLeuGluGluThrGlyGlySerLleIleGly 259
|||||
725 TTGGGAAGGTTTGGAAACTAGAGCAAAATGTTGGAAGCATATAGTGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
|||||
775 ACTTTTAAAGCAATTCAGAGAGGAAAAAGCGTCCAGAGCGAAGAGAGA 824
276 pProArgLeuProLysProLysGlyGluThrValGlySerPheArgLysG 293
|||||
825 CCGCGCGCTGCCAAAACACAGAGGCCAAACAGTTGTTCTTTCAGAGAGG 874
293 lYLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
|||||
875 GACTTCGATGTTGCCAGAGCAATATCTGCAGATTAGTGACAAAGTT 924
310 LysLeuSerTrrPlyLysSerSerLleThrLysSerGluLysGlyGly 326
|||||
925 AAGTGTCTTGGAGGCTCTCAGGTATCACTAAAGCTGAGAGCGAGAGATA 974
326 rHisLeuThrTrrGluThrProGluGlyValAlaSerLeuGlnSerArgS 343
|||||
975 CAACCTTAACATATGAGACTCCAGATGTTAGTTCTCCGACAGACAAA 1024
343 erLleValMetThrValProSerTrrValAlaSerAsnLleLeuArgPro 359
|||||
1025 GTGTGTATATGAGCGTGCATCTCATGTTGCAAGTGTCTTGCOCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTrrTrrProProVa 376
|||||
1075 CTTTTCGAATCTGCTGCAATATGACACTCTCAAAACATATATACCCACAGT 1124
376 lGlyAlaValThrLleSerTrrProGluGluAlaIleArgAspGluArgL 393
|||||
1125 TGCAGCAGATTCATCTCGTACCCGAAAGAGCAATCCAGACGAATGTT 1174
393 euValAspGlyLeuLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
|||||
1175 TGATATAGTGTGAACTAAAGGGTTTGGCAATGTCATCCAGCGACGCA 1224
410 GlyValGluThrLeuGlyThrLleTrrSerSerSerLeuPheProAsnAr 426
|||||
1225 GGGATTGAACATTAGAACATCTACACTCTCTTCCCAATG 1274
426 gAlaProLysGlyArgValLeuLeuLeuAsnTrrLleGlyValAlaLysA 443
|||||
1275 CGCACCGCCCGAGAGATTTGCTGTGAACTACATGCGGGGTCTACAA 1324
443 snProGluLleLeuSerLysThrGluSerGlnLeuValGluValAlaAsp 459
|||||
1325 ACACCGGAATTTGCTCAAGTCTGAAGTGAGTTAGTGAACAGTTGAC 1374
460 ArgAspLeuArgLysMetLeuLleLysProLysAlaGlnAspProLeuVa 476
|||||
1375 AGAGATTGAGGAAATGCTAATTAAGCTTAATGCACGATCCACTTAA 1424
476 lValGlyValArgValTrrProGluAlaIleProGlnPheLeuValGlyH 493
|||||
1425 ATTAGAGATTAGGTTAGGCTCAAGCCATTCCTAGTTCTAGTTGGTC 1474
493 lAsLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
|||||
1475 ACTTTGATATCTTGACAGCGCTAAATCATCTCTAACGCTTCCGGGCTAC 1524
510 GluGlyLeuPheLeuGlyLysAsnTrrValSerGlyValAlaLeuGlyAr 526
|||||
1525 GAAGGCGCTATTTTGGTGGCAATTAACGCTGCTGTAGCCTTAGCGCG 1574
526 gcysValGluGluAlaTrrGluValAlaSerGluValThrGlyPheLeuS 543

|||||
1575 GTGTGTAGAGCGCATTTGAAACCGCATTTGAGGTCAACACTTCACT 1624
543 eArgTrrAlaTrrLys 548
|||||
1625 CACGCTACGCTTACAG 1641
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.AAC47180
seq_documentation_block:
ID AAC47180 standard; DNA; 1690 BP.
XX
AC AAC47180;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52854.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134378.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145293.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155693.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159332.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160981.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores:

Quality: 2075.00 Length: 536
Ratio: 4.235 Caps: 4
Percent Similarity: 91.418 Percent Identity: 75.373

alignment_block:

US-09-508-418-2 x AAC47180 ..

Align seg 1/1 to: AAC47180 from: 1 to: 1690

```

22 LeuAlaIhLeuSnaArgThrSerPheIleProPheSerSerIleSerIy 38
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7  TTATCTCTCTCCGTCGACGACCAATCGCTTCTCCGCTTTCGAA 56
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
38 sArgAsnSerValAsnCysAsnGlyTrp.....ArgThrArgCys 52
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 GCCCAAT...CTCCGATTAAATGTTTAAAGCCCTTAGACTCCGTTGTT 103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
52 eValAlaIyAspTyrThrValIProSerSerAlaValAspGlyIlePro 68
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
104 CAGTGGCCGGTGGACCAACCGCTGGAATCTCAAAATCGAAGCGGAGGA 153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 AlaAlaIleu.....AspCysValIleValGlyIleSerG 83
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
154 GGCACCCACCACTCACAGGATGTGTGATGTGCGGAGGATATAGTGG 203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
83 yLeuCysIleAlaGlnValMetSerAlaAsnTyr.....ProA 96
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
204 TCTTGGATGCTCAGCGCTGCTCTACTAAGCATCCTGATGCTGCCGA 253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
96 sLeuMetValThrGluAlaArgAspArgAlaGlyIleAsnIleThrThr 112
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
254 ATTTAAATTTGACCGAGCGTAAGGATCGTGTGAGGCAACATTATCACT 303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
113 ValGluArgAspGlyTyrIleuTrpGluGlyIleProAsnSerPheGlnPr 129
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
304 CGTGAAAGAGATGTTTCTCTGGAAGAGTCCCAATAGTTTCAACC 353
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
129 oSerAspProMetIleuThrMetAlaValAspCysGlyIleuLysAspSpl 146
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
354 GTCTGATCTATGCTCACTATGCTGAGTATGTTGTAAGATGATTT 403
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
146 euValLeuGlyAspProAsnAlaProArgPheValIleuTrpLysGlyLys 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
404 TGGTGTGGAGATCCTACTGCCCAAGTTTGTGTGTGATGGGAAA 453
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
163 LeuArgProValProSerLysLeuThrAspLeuProPhePheAspLeuKe 179
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
454 TTGAGCGCGGTTCCATCGAAGCTAACAGACTTACCGTTTGTATTTGAT 503
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
179 tSerIleProGlyLysLeuArgAlaGlyIlePheGlyProIleGlyLeuArgP 196
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
504 GAGTATTGTTGGGAGATTAAGACTGTTTGGTGCACCTTGGCATTCGAC 553
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
196 roSerProGlyHisGluGluSerValGluGlnPheValArgAspAsn 212
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
554 CGTCACCTCCAGGTCGTGAAGAAATCTGTGAGGAGATTGTACGGCGTAAC 603
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
213 LeuGlyGlyIleValPheGluArgLeuIleGluProPheCysSerGlyVa 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
604 CTGCGTATGAGGTTTGTAGCGCTGATGTAACCGTTTGTTCAGGTGT 653
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
229 lTyrValGlyAspProSerLysLeuSerMetLysAlaIlePheGlyLysV 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
654 TTTATGCTGGGATCTTCAAAACCTGAGCAAGAAAGCAGCGTTGGGAAG 703
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
246 alTrpLysLeuGluGlnThrGlyIleSerIleIleGlyIleThrPheLys 262
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
704 TTTGGAACCTAGAGCAAAATGTTGGAAGCATTAATAGTGTACTTTTAA 753
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
263 AlaIleLysGluArgSerSerThrProLysAlaProArgAspProArgLe 279
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
754 GCAATTCAGAGAGGAAAGCGCTCCCAAGCGAAGCAAGACCGCGGCT 803
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
279 uProLysProLysGlyIleThrValGlySerPheArgLysGlyLeuArgm 296
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
804 GCCAAAACCAAGGCGCAACAGTGTCTTTCAGGAAGGCGACTTCGAA 853
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
296 etcLeuProAspAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSer 312
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
854 TGTTCGCAAGAGCAATATCTGCAAGATTAGTACAAAGTTAAGTTGTCT 903

```

```

313 TrpLysLeuSerSerIleThrLysSerGluIyGlyIleThrHisLeuTh 329
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
904 TGAAGCTCTCAGTATCACTAAGCTGGAGAGCGGAGATCAACTTAAC 953
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
329 rTyrGluThrProGluGlyValValSerLeuGlnSerArgSerIleValm 346
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
954 ATATAGACTCCAGATGTTTGTTCGTCGAGAGCAAAAGTGTGTAA 1003
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
346 eThrValProSerTyrValAlaSerAsnIleLeuArgProIleuSerVal 362
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1004 TCACGCTGCATCTATATGTGCAAGTGTCTTCCGCCCTTTCTGAA 1053
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
363 AlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaVa 379
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1054 TCTGCTGCAAAATGACTCTCAAAACTATATTATCCACAGATTGCAGCAGT 1103
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
379 lThrIleSerTyrProGluGlnAlaIleArgAspGluArgLeuValAspG 396
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1104 ATCTATCTCTGACCCGAAAGAAAGCAATCCGAACAGAAAGTTTGATAG 1153
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
396 lYgluLeuLysGlyPheGlyGlnLeuHisProArgThrGlnIyValGlu 412
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1154 GTGACTTAAGGCTTTTGGGCAATTGCATCCACGACCGCAAGAGATTGAA 1203
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
413 ThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLy 429
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
429 sGlyArgValLeuLeuLeuAsnTyrIleGlyIleAlaLysAsnProGluI 446
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1254 CGGAAGAAATTTGCTGTGAACCTATACATTGGCGGGTCTCAAAACACCG 1303
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
446 lLeuSerLysThrGluSerGlnLeuValGluValAlaAspArgAspLeu 462
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1304 TTCTGTCCAGCTCTAAGGTGAGTGAAGTGAAGCAAGCTTGACAGAGATTG 1353
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
463 ArgLysMetLeuIleLysProLysAlaGlnAspProLeuValValGlyVa 479
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1354 AGGAAATCTTAATTAAAGCTAATTCGACCGCATCCACTTAATTAGAGAT 1403
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
479 lArgValITrPProGlnAlaIleProGlnPheLeuValLysHisLeuAspT 496
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1404 TTAGGTATAGGCTCAAGCATCTCCAGTTTCTAAGTGTGCTACATTGATA 1453
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
496 hTrLeuSerThrAlaLysAlaIleMetAsnAspAsnGlyLeuGluIleu 512
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1454 TCTTGACACGCGCTAATCATCTTAACGCTCTTCGGGCTACGAAGGCTA 1503
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
513 PheLeuGlyGlyAsnTyrValSerGlyValAlaIleLeuGlyArgCysValG 529
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1504 TTTTGGGGGCAATTACGTCGCTGTGTGAGCCTTAAGCCGCGTGTGTGA 1553
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
529 uGlyAlaTyrGluValAlaSerGluValIThrGlyPheLeuSerArgTyrA 546
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1554 AGCGCATATGAACCGGATGAGGTCAACAACCTTATGTCACGGTAGC 1603
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
546 lATrLys 548
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1604 CTTACAG 1611
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF76571
seq_documentation_block:
ID AAF76571 standard; cdNA; 1719 bp.
XX
XX AAF76571;
AC
XX
XX
DT 16-MAY-2001 (first entry)
DE Arabidopsis protoporphyrinogen oxidase coding sequence SEQ ID NO: 1.
XX
KW Protoporphyrinogen oxidase; protox; herbicide-tolerance; wheat; rice;

```



```

426 gAlaProlySGlyArgValLeuLeuLeuAsnTyrIleGlyAlaAlaLysA 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1275 CGACCGCCGCGGAGAAATTTGCTGTGAACCTACATTGGCGGCTACAA 1324
443 snProGluIleLeuSerLysThrGluSerGluLeuValGluValAsp 459
    || ||||| ||||| ||||| ||||| ||||| ||||| |||||
1335 ACACCGGAATTTCTCTCAAGCTCAAGCTCAAGTGAAGGAAAGCTTGAC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1375 AGACATTGAGGAAATGCTAATTAAGCTTAATTCGACCATCAGCTTAA 1424
476 lValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGlyH 493
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1425 ATTAAGGAGTAAAGGATAGGCTCAAGCCATTCCTCAAGTTCTAGTGGTC 1474
493 lIleAspTrpLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1475 ACTTTGATATCCTTTGACACGCGCTAAATCATCTCAACGCTCTGGGCTAC 1524
510 GluGlyLeuPheLeuGlyGlyAsnTyr-ValSerGlyValAlaLeuGlyAr 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1525 GAAGGCTATTTTGGTGCGCAATTAGCTGCGTGTCTAGCCTTAGGCCG 1574
526 GCGValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1575 GTGTGTGAAGGCGGCATATGAACCGCATTTGAGTGCAACACTCATGT 1624
543 eArgTyrAlaTyrLys 548
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1625 CACGCTACGCTTACAG 1641

seq_name: /SID8/gc9data/geneseq/geneseqn/NA2000.DAT:AA245271
seq_documentation_block:
ID AA245271 standard: DNA; 1632 BP.
AC AA245271;
AT 27-MAR-2000 (first entry)
DE DNA encoding the chloroplast-type soybean protoporphyrin IX oxidase.
KW Weed control; weed activity; plant; protoporphyrin IX oxidase; PPO;
    herbicide; chloromethoxymyl; acifluorfen; protoporphyrin X; soybean; ss.
OS Glycine max.
XX
XX
XX Key Location/Qualifiers
    CDS 1..1632
        /tag= a
        /product= "protoporphyrin IX oxidase"
XX
XX EP953646-A2.
XX
XX PD 03-NOV-1999.
XX
XX PF 30-APR-1999; 99EP-0108463.
XX
XX PR 30-APR-1998; 98JP-0120553.
XX PR 02-OCT-1998; 98JP-0281127.
XX PR 20-NOV-1998; 98JP-0330981.
XX PR 02-MAR-1999; 99JP-0054730.
XX
XX (SDMO ) SUMITOMO CHEM CO LTD.
XX
XX PA Nakejima H, Nagasawa A.
XX
XX PI WPL: 2000-088762/08.
XX
XX DR P-PSDB: AA245100.
XX
XX PT Method of making plants resistant to weed control compounds by

```

```

PT introducing gene encoding protein into plant cell -
XX
XX PS Example 3; Page 60-64; 119pp; English.
XX
XX CC The present sequence encodes the soybean protoporphyrin IX oxidase
    CC (PPO) gene. The PPO gene and protein are used in the course of the
    CC invention. The specification describes a method of making plants
    CC resistant to weed control compounds. The method comprises introducing
    CC a gene encoding a protein into a plant cell, where the protein has a
    CC specific affinity for a substance which is concerned with the weed
    CC activity of a weed control compound, e.g. protoporphyrin IX; has no
    CC ability to modify this protein; and is virtually free from framework
    CC regions in an immunoglobulin; and expressing the gene. The gene of the
    CC invention is used for producing a plant resistant to specified compounds,
    CC such as protoporphyrin IX oxidase (PPO) inhibitory-type herbicides, e.g.
    CC chloromethoxymyl, acifluorfen, etc. When PPO is inhibited,
    CC protoporphyrin IX (the substrate of PPO) accumulates in the plant cells,
    CC and is metabolised to form protoporphyrin X, followed by formation of
    CC active oxygen in the presence of protoporphyrin X and light, which
    CC damages cell functions.
XX
XX SQ Sequence 1632 BP; 404 A; 387 C; 409 G; 432 T; 0 other;

```

```

alignment_scores:
    Quality: 2062.50      Length: 550
    Ratio: 4.218          Gaps: 7
    Percent Similarity: 88.909    Percent Identity: 73.273

```

alignment_block:

US-09-508-418-2 x AA245271 ..

Align seg 1/1 to: AA245271 from: 1 to: 1632

```

3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
    ::::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
48 TCTTCCGCCCTCTCCATTCCTCCCAACCTTCTCTCACTCCCTCCACTC 97
19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerI 36
    ||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
98 GAAA.....ATTCCTC.....T 108
36 lE-SerLysArgAsnSerValAsnCysAsnGlyTrpArgTrpArgCysSe 52
    ||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
109 CCCTCTCCGCCCTTAACCTAT.....CTACGCTGCTC 140
52 rValAlaLysAspTyrThrVal...ProSerSerAlaValSpGlyGlyP 68
    ||::: ||||| ||||| ||||| ||||| ||||| ||||| |||||
141 CATTCGGGAGGAATCCACCGCTCTCCGCCCAAAACGACAGACTCCGCC 190
68 rAlaAlaGluLeuAspCysValIleValGlyAlaGlyTlSerGlyLeu 84
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
191 C.....GTGACATCGCTGCTGCTGCGGAGGAGCTTACGCGCTC 231
85 CysIleAlaGluValMetSerAlaAsnTyr.....ProAsnLeuMetVa 99
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 TGCATCGCCCGAGCCCTCGCCCAACGCAATGCCAAGCTGCTGT 281
99 lThrGluAlaArgAspArgAlaGlyGlyAsnIlePheThrValGluArgA 116
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 CACGAGGCGCCGAGACCGGCTGCGGCGCAACATCCACGATGAGGAGG 331
116 spGlyTyrLeuTrpGluGlyProAsnSerPheGlnProSerAspPro 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 ACGGATACCTCTGGGAGAAGGCCCAACAGCTTCCACCTTGATGATCA 381
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspLeuValLeuGl 149
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
382 ATGCTACACCATGCTGTGACAGTGTAAAGATGAGTGTGTTGGG 431
149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgPro 166
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
432 GATCTCTATGACCTCGGTTTGTGTGAAACAGAAATTGAGGCGCG 481

```

```

166 aProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 TGCCGGGAGAGCTGACTGATTTGCCCTTCTTTCGACTGATGAGATTGGT 531
183 GlyLysLeuAArgAlaGlyPheGlyProIleGlyLeuAArgProSerProP 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
532 GGCAAAATACAGGGCGGTGGCTTGGTGGCGTTGGAAATGGCGCTCCTCC 581
199 OGlyHISGluGluSerValGluGlnPheValAArgAsnLeuGlyG 216
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
582 AGGTCATGAGAAATCGGTTGAAGAGTTGTCGCGGAACCTTGCTGATG 631
216 LuvalPheGluAArgLeuIleGluProPheCysSerGlyValTyrValGly 232
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
632 AGGTTTGAACGGTTGATAGAGCCTTTGTTGTCAGGGGCTATAGCAGGC 681
233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTyrPlySle 249
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
682 GATCCTTCAAAATTAGATAGAAAGCAGACATTCGGAAAGTTGGAAGCT 731
249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
732 GCAAAAATATGCTGTAGCATTTATGCGAACTTCAAAACCAATACAG 781
266 LuArgSerSerThrProLysAlaProAArgAspProArgLeuProLysPro 282
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
782 AGAGAAATGAGAGCTTCANAAACCACCTCGAGATCCGCGCTCCCAAAACA 831
283 LysGlyGlnThrValGlySerPheAArgLysGlyLeuAArgMetLeuProAs 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
832 AAAGGTCAGAGCTGTGATCTTTCGGAAGGAGACTTACCATGTTGCCGGA 881
299 PAlaIleSerAlaAArgLeuGlySerLysLeuLysLeuSerTrpLysLeus 316
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
882 TGCATTTTCTGCCAGACTAGCGCAACAAAGTAAGTTATCTGGAAAGCTTT 931
316 eSerIleThrLysSerGluLysGlyGlyTyrHISLeuThrTyrGluThr 332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
932 CAAGATATAGTAACACTGATGAGAGTAGACAGTTGACATATGAAACA 981
333 ProGluGlyValValSerLeuGlnSerAArgSerIleValMetThrValPr 349
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
982 CCAGAGAGAGTGGTTCTTTCAGTCAAAACCTGTTCTCGACCATTTCC 1031
349 oSerTyrValAlaSerAsnIleLeuAArgProLeuSerValAlaAlaAla 366
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1032 TTCTATGTGTGCTAGTACATGCTGCGTCCTGCTGCTGCTGCTGCGAG 1081
366 sPAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1082 ATGCACCTTCAAAAGTTTATACCTCCACGTTGCTGCGAGTTCCATATCC 1131
383 TyrProGlnGluAlaIleAArgAspGluAArgLeuValAspGlyGluLeu 399
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1132 TATCCAAAAGAGCTATTAGATACAGAAATGCTTAGATGCTGAGTTGAA 1181
399 sGlyPheGlyGlnLeuHISProAArgThrGlnGlyValGluThrLeuGlyT 416
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1182 GGGGTTTGGTCATATGATCCACGTAAGCAAGGAGTGAACATTAGGAA 1231
416 hIleTyrSerSerSerLeuPheProAsnAArgAlaProLysGlyAArgVal 432
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1232 CTATATACAGCTATCATTATCCCAACGAGACACACCTCGGAAGGTT 1281
433 LeuLeuLeuAsnTyrIleGlyValAlaLysAsnProGluIleLeuSerLys 449
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1282 CTACTCTTGAATTACATGAGAGCACTAATCTGGAATTTATATCGAA 1331
449 sThrLysGlnLeuValGluValValAspAArgAspLeuArgLysMetL 466
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1332 GACGGACAGTGAACCTTGTGAAACAGTTGATCGAGATTGTGAGAAATCC 1381

```

```

466 euIleLysProLysAlaGlnAspProLeuValValGlyValAArgValTrp 482
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1382 TTTATACCCAAATGCCAGATCCATTTGTAGTGGGGGTGAGACTGGG 1431
483 ProGlnAlaIleProGlnPheLeuValGlyHISLeuAspThrLeuSerTh 499
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1432 CCTCAGACTTATCCACAGTTCTTAGTTGGCCATTCGATCTCTAGATCT 1481
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1482 TGCTAAAGCTTCTATCAGAAATAGTGGTTGAAAGGCGCTTCTTGGGG 1531
516 LyAsnTyrValSerGlyValAlaLeuGlyAArgCysValGluGlyAArg 532
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1532 GTATATTATGTTCTGTGTTGCCCTTGAGACGATGCTTCAGGAGAGCTAT 1581
533 GluValAlaSerGluValThrGlyPheLeuSerAArgTyrAlaTyrLys 548
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1582 GAGGTAGCAGCTGAGTAACGATTTTCTCACAAAATAGAGTGTACAAA 1629
seq_name: /SIDS8/gcdata/geneseq/geneseqn/NA197.DAT:AAV04309
seq_documentation_block:
ID AAV04309 standard; cDNA; 1847 BP.
XX
XX AAV04309;
XX
XX 20-APR-1998 (first entry)
XX
XX Soybean protox-1 cDNA.
XX
XX Protoporphyriinogen oxidase-1; protox-1; promoter; soybean;
XX herbicide resistance; breeding programme; probe; gene isolation;
XX genomic mapping; ss.
XX
XX Glycine max.
XX
XX OS
XX FH Key Location/Qualifiers
XX FT CDS 55..1686
XX FT /tag= a
XX FT /product= protox-1
XX
XX W09732028-AL.
XX
XX PD 04-SEP-1997.
XX
XX PF 27-FEB-1997; 97WO-US03343.
XX
XX PR 21-JUN-1996; 96US-0020003.
XX PR 28-FEB-1996; 96US-0012705.
XX PR 28-FEB-1996; 96US-0013612.
XX
XX PA (NOVS ) NOVARTIS AG.
XX
XX Johnson MA, Volrath SL, Ward ER:
XX
XX MPI: 1997-489209/45.
XX
XX DR P-PSDB; AAW41608.
XX
XX PT DNA containing a plant proto-porphyrinogen oxidase gene promoter -
XX optionally linked to a heterologous gene, especially to express
XX herbicide-resistant enzymes, and plants containing such constructs
XX
XX PS Disclosure: Pages 68-71, 114pp; English.
XX
XX CC The present sequence encodes soybean protoporphyriinogen oxidase-1
XX (protox-1).
XX CC The protox-1 promoter can be used to express herbicide resistant
XX CC enzymes, specifically protox, i.e. a plant tissue, plant or progeny
XX CC containing a chimeric gene of the promoter and a heterologous
XX CC coding sequence. The plant can also be used in breeding programmes.
XX CC Also hybridising fragments of the protox coding sequence can be
XX CC used as probes, e.g. to isolate related genes or for genomic

```

CC mapping.
 XX Sequence 1847 BP; 473 A; 416 C; 453 G; 505 T; 0 other;
 alignment_scores: Quality: 2062.50 Length: 550
 Ratio: 4.218 Gaps: 7
 Percent Similarity: 88.909 Percent Identity: 73.273

alignment block:
 US-09-508-418-2 x AAV04309 ..

Align seg 1/1 to: AAV04309 from: 1 to: 1847

```

3 ThrThrProIleAlaIAsnHisProAsnIlePheThrHisGlnSerSerse 19
   :::::|||||:  |||:|||||:  |||  |||  |||
102 TCCTCCGCCCTCCCTCCATTCGCCCAACCTTTCTTCACCTCCACACTC 151
   :::::|||||:  |||:|||||:  |||  |||  |||
19 rSerProIleuAlaPheLeuAsnArgThrSerPheIleProPheSerSer 36
   :::::|||||:  |||:|||||:  |||  |||  |||
152 GAAA.....ATTTCC.....T 162
   :::::|||||:  |||:|||||:  |||  |||  |||
36 le.SerIlyAsnArgSernSerValAsnCysAsnGlyTTPArgThrArgCysse 52
   :::::|||||:  |||:|||||:  |||  |||  |||
163 CGCTCTCGCCCTAACCTTAFT.....ATTCCGCTGCTC 194
   :::::|||||:  |||:|||||:  |||  |||  |||
52 rValAlaIAsnAspTyrThrVal...ProSerSerAlaValAsnGlyGlyP 68
   :::::|||||:  |||:|||||:  |||  |||  |||
195 CATTGGCGAGAGATCCACCGCGTCTCCGCCCAAAACACAGACTCCGCC 244
   :::::|||||:  |||:|||||:  |||  |||  |||
68 roAlaIaIaGluIeuAspCysValIleValIaGlyAlaGlyIleSerGlyLeu 84
   :::::|||||:  |||:|||||:  |||  |||  |||
245 CC.....GTGACTGCGTGTGCTGCGCGAGGCGCTACAGCGCCCTC 285
   :::::|||||:  |||:|||||:  |||  |||  |||
85 CysIleAlaGlnValMetSerAlaAsnTyr.....ProAsnLeuMetVal 99
   :::::|||||:  |||:|||||:  |||  |||  |||
286 TGGATGCGCCGAGCCCTCGCCACCAACAGCCCAATGCCAACGTCGCTGT 335
   :::::|||||:  |||:|||||:  |||  |||  |||
99 lThrGlnAlaIaArgAspArgAlaGlyIleAsnIleThrThrValIaIaArg 116
   :::::|||||:  |||:|||||:  |||  |||  |||
336 CACGGAGCGCCGAGACCGCGTGTGCGCGCAACATCACACGATGAGAGGG 385
   :::::|||||:  |||:|||||:  |||  |||  |||
116 spGlyTyrIleuTrpGlnGluGlyProAsnSerPheGlnProSerAspPro 132
   :::::|||||:  |||:|||||:  |||  |||  |||
386 ACCGATATCTCTGTGGAAAGAGGCCCAACAGCTTCCAGCCTTGTATCCA 435
   :::::|||||:  |||:|||||:  |||  |||  |||
133 MetIleuThrMetAlaValAspCysGlyLeuIlyAspAspLeuValIeugl 149
   :::::|||||:  |||:|||||:  |||  |||  |||
436 ATGCTCACCACTGGTGTGGACAGTGTAAAGATGAGACTGTGTTTGGG 485
   :::::|||||:  |||:|||||:  |||  |||  |||
149 yAspProAsnAlaProArgPheValIleuTrpLysGlyIlyLeuArgProV 166
   :::::|||||:  |||:|||||:  |||  |||  |||
486 GGAATCCCGATGACACCTCGGTTGTGTGTGGAAACAGAACTTGAGCCGG 535
   :::::|||||:  |||:|||||:  |||  |||  |||
166 alProSerIlyLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
   :::::|||||:  |||:|||||:  |||  |||  |||
536 TGCCCGGAGAGCTGACTGATTTGCCTTTCTTGACTTGATGACATTTGGT 585
   :::::|||||:  |||:|||||:  |||  |||  |||
183 GILyIleuAsnArgAlaGlyPheGlyProIleGlyIleuAsnArgProSerPro 199
   :::::|||||:  |||:|||||:  |||  |||  |||
586 GGCAGAAATCGGGCTGTGGTGTGGCTGTGGCAATTCGGCTCTCTCC 635
   :::::|||||:  |||:|||||:  |||  |||  |||
199 ogIlyHisGlnGluSerValGlnGlnPheValaIaArgAsnLeuGlyGly 216
   :::::|||||:  |||:|||||:  |||  |||  |||
636 AGGTCAATGAGGATCGGTGAAGATTGTGTCTCGGAACCTTGCGTGAAG 685
   :::::|||||:  |||:|||||:  |||  |||  |||
216 luValPheGluArgLeuIleGlnProPheCysSerGlyValIyValIy 232
   :::::|||||:  |||:|||||:  |||  |||  |||
686 AGGTTTTTGAACGTGTATAGACCTTTTGTTCAGGGGTCTATCAGAGCC 735
   :::::|||||:  |||:|||||:  |||  |||  |||
233 AspProSerIlyLeuSerMetIlyAsnAlaIaPheGlyIlyValTrpLysIe 249
   :::::|||||:  |||:|||||:  |||  |||  |||

```

```

736 GATCCTTCAAAATTAAGTATGAAAGCACATTCGGGAAAGTTTGAACCT 785
249 uGluGluThrGlnGlySerIleIleGlyIleThrPheIlyAlaIleIySG 266
   :::::|||||:  |||:|||||:  |||  |||  |||
786 GGAAGAAATGTGTGTAGCATTTATTTGTGGAACCTTCAACAGCATACAG 835
   :::::|||||:  |||:|||||:  |||  |||  |||
266 luArgSerSerThrProIlyAsnAlaProArgAspProArgLeuProIlyPro 282
   :::::|||||:  |||:|||||:  |||  |||  |||
836 ACAGAAATGAGACTTCAAACACACCTCGAGATCCGCGTCTGCCAAACCA 885
   :::::|||||:  |||:|||||:  |||  |||  |||
283 LysGlyGlnThrValIleGlySerPheArgIySGIlyLeuArgMetLeuProAs 299
   :::::|||||:  |||:|||||:  |||  |||  |||
886 AAAGCTCAGACTGTGTGATCTTCCGGAAGGACTTACCATGTGTCCGCA 935
   :::::|||||:  |||:|||||:  |||  |||  |||
299 palAlleSerAlaArgLeuGlySerIlyLeuIlySerIlyLeuSerTrpLysLeu 316
   :::::|||||:  |||:|||||:  |||  |||  |||
936 TGCAATTTCTGCCAGACTTAGGACACAAAGTAAAGTATCTTGGAAAGCTT 985
   :::::|||||:  |||:|||||:  |||  |||  |||
316 erSerIleThrIlySerSerGluIySGIlyTyrHisLeuThrTyrGluThr 332
   :::::|||||:  |||:|||||:  |||  |||  |||
986 CAAGTATTAGTAACTGGATAGTGAGAGTACAGTTGACATATGAAACA 1035
   :::::|||||:  |||:|||||:  |||  |||  |||
333 ProGluGlyValaIleSerLeuGlnSerArgSerIleValMetThrValPr 349
   :::::|||||:  |||:|||||:  |||  |||  |||
1036 CCAGAGAGAGTGTCTTGTGCAAGTGAACAACTGTGTCTGACCATCTCC 1085
   :::::|||||:  |||:|||||:  |||  |||  |||
349 oSerTyrAlaIaSerAsnIleLeuArgProLeuSerValAlaIaIaIa 366
   :::::|||||:  |||:|||||:  |||  |||  |||
1086 TTCCTATGTGTGTATACATTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1135
   :::::|||||:  |||:|||||:  |||  |||  |||
366 spAlaIleuSerAsnPheTyrTyrProProValaIleValaIaIaIleSer 382
   :::::|||||:  |||:|||||:  |||  |||  |||
1136 ANGCACTTCAAAAGTTTATATACCTCCAGTGTGCTGCAAGTTCCATATCC 1185
   :::::|||||:  |||:|||||:  |||  |||  |||
383 TyrProGlnGluAlaIleArgAspGluArgLeuValaIaIaGlyIleuIly 399
   :::::|||||:  |||:|||||:  |||  |||  |||
1186 TATCCAAAAGAGCTTATAGACGAATGCTGATGATGTGTGAGTTGAA 1235
   :::::|||||:  |||:|||||:  |||  |||  |||
399 sGlyPheGlyGlnIleuHisProArgThrGlnGlyValaIaIaIleuGlyT 416
   :::::|||||:  |||:|||||:  |||  |||  |||
1236 GGGGTTGTGTCATTTGATCCACGTCAGCCAAAGAGTGAACATTTAGAA 1285
   :::::|||||:  |||:|||||:  |||  |||  |||
416 hIleTyrSerSerSerLeuPheProAsnArgAlaIaIaIySGIlyValaI 432
   :::::|||||:  |||:|||||:  |||  |||  |||
1286 CTATATACAGCTCATCACTATTCGCCAACGAGCCACCTGGAAGGGTT 1335
   :::::|||||:  |||:|||||:  |||  |||  |||
433 LeuLeuLeuAsnTyrIleGlyIlyAlaIyAsnProGluIleIleuSerIly 449
   :::::|||||:  |||:|||||:  |||  |||  |||
1336 CTACTCTGAATTTACATTGGAGAGCAACTAATCTGGAATTTTATCGAA 1385
   :::::|||||:  |||:|||||:  |||  |||  |||
449 sThrGluSerGlnIleuValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 466
   :::::|||||:  |||:|||||:  |||  |||  |||
1386 GAGCGACAGTGAACCTTGTGGAAACAGTTGATCGAGATTTGAGGAAATCC 1435
   :::::|||||:  |||:|||||:  |||  |||  |||
466 euIleIySProlYsAlaGlnAspProLeuValaIaIyValaIaIyVal 482
   :::::|||||:  |||:|||||:  |||  |||  |||
1436 TTATATAACCCAAATCCAGATCCATTTGTAGTGGGGGTGAGACTGTGG 1485
   :::::|||||:  |||:|||||:  |||  |||  |||
483 ProGlnAlaIleProGlnPheLeuValaIaIyHisLeuAspTrpIleuSerTh 499
   :::::|||||:  |||:|||||:  |||  |||  |||
1486 CCTCAAGCATTTCCACAGTTCTTAACTGGCCATCTGTGATCTTCTAGAGT 1535
   :::::|||||:  |||:|||||:  |||  |||  |||
499 rAlaIySAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 516
   :::::|||||:  |||:|||||:  |||  |||  |||
1536 TGCTTAAGCTTCTATACAAATACGTGGTGTGAAGGGCTCTCTCTGGGG 1585
   :::::|||||:  |||:|||||:  |||  |||  |||
516 lyAsnTyrValSerGlyValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 532
   :::::|||||:  |||:|||||:  |||  |||  |||
1586 GTAATATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1635
   :::::|||||:  |||:|||||:  |||  |||  |||
533 GluValaIaIaSerGluValaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 548
   :::::|||||:  |||:|||||:  |||  |||  |||
1636 GAGTAGCAGCTGAAGTAAACGATTTTCTCAAAATAGAGTGAACAA 1683
   :::::|||||:  |||:|||||:  |||  |||  |||

```



```
336 CACGGAGGCCGAGACCGCGTCGGCGGCAACATCACCAAGATGGAGAGCG 385
116 SPGLYTYRLeuTrpGluGluGluProAsnSerPheGlnProSerAspPro 132
386 ACGGATACCTCTGGGAGAGAGGCCCAACAGCTTCCAGCCTTCGATGCCA 435
133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGlu 149
436 ATGCTCAACCATGTGTGGAGACAGTGTAAAGATGAGCTTGTGTTGGG 485
149 YAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArgProV 166
486 GGATCTGATGACCTCGGTTGTGTGTGGACAGGAGATTGGAGCCGG 535
166 aProSerLysLeuThrAspLeuProPheAspLeuMetSerIlePro 182
536 TGCCCGGAGACCTGACTGATTTGCCCTTCTTGACTTGATGAGCATTTG 585
183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSerProP 199
586 GGCATAAATCAGGGCTGGCTTGTGGCTGGAAATTCGGCTCCTCCTCC 635
199 OGlyHisGluGluSerValGluGlnPheValArgArgAsnLeuGlyGly 216
636 AGGTATATAGCAATCGTTGAAGAGTTGTTCTCGGGAACCTTGATG 685
216 LuValPheGluLeuArgLeuIleGluProPheCysSerGlyValTyrValGly 232
686 AGGTTTGTGAACGGTTGATAGAGCCTTTTGTTCAGGGGCTATAGCAGGC 735
233 AspProSerLysLeuSerMetLysAlaIlePheGlyLysValTyrLysLe 249
736 GATCCTTGAATTAAGTATGAAAGCAGCATTCGGAAAGTTTGGAAAGCT 785
249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
786 GGAATAAATGGTGGTACATATATGTCGAACCTTCAAGCAATACAAAG 835
266 LuArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282
836 AGAGAAATGAGACTTCAAAACCCACTCGAGATCCGCTGCCAAACCA 885
283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
886 AAAGGTACAGACTGTGGATCTTCCGGAAGGACTTACCATGTTGCCCTGA 935
299 PAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTyrLysLeus 316
936 TGCATATTTCTGCCAGACTAGGCACAAAGTAAAGTATCTTGAAGCTTT 985
316 eTSerIleThrLysSerGluLysGlyTyrHisLeuThrTyrGluThr 332
986 CAAGTATTAGTAAGTGTAGTGAAGTACAGTTTGAACATATGAACA 1035
332 eTSerIleThrLysSerGluLysGlyTyrHisLeuThrTyrGluThr 332
1036 CCAGAAAGAGGTGTTCTTTCGAGTGAACAAACTGTTGCTGCAGCATTC 1085
333 ProGluGlyValIleValSerLeuGlnSerArgSerIleValMetThrValPr 349
1086 TTCCTATGTTCAGTACATGCTGCTGCTCTGCTGCTGCTGCTGCTGCTG 1135
349 oSerTyrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
1136 ATGCACTTTCAAAGTTTATATACCTCCAGTGGTGCAGATTCCATATCC 1185
366 sPAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThrIleSer 382
1186 ATGCACTTTCAAAGTTTATATACCTCCAGTGGTGCAGATTCCATATCC 1185
383 TyrProGlnGluAlaIleArgAspGluArgLeuValAspGlyGluLeuLys 399
1186 TATCCAAAGAGCTATTTGATCAGATGCTTGAATGCTGAGACTTGA 1235
399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeuGlyT 416
1236 GGGGTTTGTGTCATTCATCCAGCTAGCCAAAGAGTGGAAACATTAAGAA 1285
416 hTrIleTyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
1286 CTATATACAGCTCATCATCTATTCGCCCAACGAGCACACACTGGAGGGTT 1335
433 LeuLeuLeuAsnTyrTrIleGlyGlyAlaLysAsnProGluIleLeuSerLy 449
1336 CTACTCTTGAAATTACATTGGAGGAGCACTAATACGGAAATTTTATCGAA 1385
449 sThrGluSerGlnLeuValGluValAlaValAspArgAspLeuArgLysMetL 466
1386 GACGACAGTGAACCTGTGGAACAGTGTGATGACAGATTGGAGAAATCC 1435
466 euIleLysProLysAlaGlnAspProLeuValGlyValArgValTyrP 482
1436 TTATTAACCCAAATGCCAGGATCCATTGTGTGTGGGGGTGAGACTGTG 1485
483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
1486 CCTCAAGCTATTCACAGCTTCTTACTGTCGCATCTTGATCTTGATGATG 1535
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeuGlyG 516
1536 TGCTAAAGCTTCTATCAGAAATCTGCTTGGAAAGGCTCTTCTTGGGG 1585
516 LysAsnTyrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
1586 GTAATTATGTGTGGTGTGTGCTTGGAGATGCTTGAAGGAGGCTAT 1635
533 GluValAlaSerGluValThrGlyPheLeuSerArgTyrAlaTyrLys 548
1636 GAGGTGACAGCTGAGTAAACGATTTTCTCACAAATAGAGTGAACAA 1683
```

OM of: US-09-508-418-2 to: Issued_Patents_NA:* out_format : pfs
Date: Jul 3, 2001 11:38 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame-pn.model -DEV=xip
-Q/cgn2_1/USPTO_spoil/US09508418/runat_03072001_100017_24452/app-query.fasta_1.613
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -YCAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blissum62
-TRAN=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM-ext -MINLEN=0 -MAXLEN=200000000
-USER=US09508418_ecgn1_1.66 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPTX
-WAIT -THREADS=1

Search information block:

Query: US-09-508-418-2
Query length: 548
Database: Issued_Patents_NA:*
Database sequences: 317530
Database length: 92630169
Search time (sec): 94.000000

score_list:

| Sequence | Strid | Orig | ZScore | BScore | Len | ! Documentation | |
|---|-------|------|---------|---------|----------|-----------------|--|
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-15 | | | 2113.00 | 4053.03 | 3.8e-218 | 1826 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-15 | | | 2113.00 | 4053.03 | 3.8e-218 | 1826 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-15 | | | 2113.00 | 4053.03 | 3.8e-218 | 1826 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-15 | | | 2113.00 | 4053.03 | 3.8e-218 | 1826 | |
| /cgn2_6/ptodata/1/ina/5A.COMB.seq:US-08-472-028A-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-09-071-296-1 | | | 2077.00 | 3984.07 | 2.6e-214 | 1719 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-11 | | | 2062.50 | 3955.03 | 1.1e-212 | 1847 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-11 | | | 2062.50 | 3955.03 | 1.1e-212 | 1847 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-11 | | | 2062.50 | 3955.03 | 1.1e-212 | 1847 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-19 | | | 2065.50 | 3846.99 | 1.1e-206 | 1784 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-19 | | | 2065.50 | 3846.99 | 1.1e-206 | 1784 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-19 | | | 2065.50 | 3846.99 | 1.1e-206 | 1784 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-931-9 | | | 1987.00 | 3809.01 | 1.5e-204 | 1811 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-9 | | | 1987.00 | 3809.01 | 1.5e-204 | 1811 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-9 | | | 1987.00 | 3809.01 | 1.5e-204 | 1811 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-9 | | | 1987.00 | 3809.01 | 1.5e-204 | 1811 | |
| /cgn2_6/ptodata/1/ina/5A.COMB.seq:US-08-472-028A-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-071-296-5 | | | 1978.00 | 3792.47 | 1.2e-203 | 1691 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-17 | | | 1933.50 | 3704.65 | 9.6e-199 | 1910 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-17 | | | 1933.50 | 3704.65 | 9.6e-199 | 1910 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-17 | | | 1933.50 | 3704.65 | 9.6e-199 | 1910 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-17 | | | 1933.50 | 3704.65 | 9.6e-199 | 1910 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-23 | | | 1823.00 | 3492.97 | 5.9e-187 | 1590 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-23 | | | 1823.00 | 3492.97 | 5.9e-187 | 1590 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-23 | | | 1823.00 | 3492.97 | 5.9e-187 | 1590 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-23 | | | 1823.00 | 3492.97 | 5.9e-187 | 1590 | |
| /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-21 | | | 1261.00 | 2407.52 | 1.7e-126 | 1224 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-08-808-323-21 | | | 1261.00 | 2407.52 | 1.7e-126 | 1224 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-050-603A-21 | | | 1261.00 | 2407.52 | 1.7e-126 | 1224 | |
| /cgn2_6/ptodata/1/ina/6A.COMB.seq:US-09-102-420B-21 | | | 1261.00 | 2407.52 | 1.7e-126 | 1224 | |

```
seq_name: /cgn2_6/ptodata/1/ina/5B.COMB.seq:US-08-808-931-15
seq_documentation block:
: Sequence 15 Application US/0808931
: Patent No. 5939602
: GENERAL INFORMATION:
: APPLICANT: Voltrath, Sandra
: APPLICANT: Johnson, Marie
: APPLICANT: Potter, Sharon
: APPLICANT: Ward, Eric
: APPLICANT: Helfetz, Peter
: TITLE OF INVENTION: DNA Molecules Encoding Plant
: TITLE OF INVENTION: Protoporphylinogen Oxidase and Inhibitor-Resistant Mutants
: TITLE OF INVENTION: Thereof
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESSES:
: ADDRESS: No. 5939602artis Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/808,931
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1847
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 15:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1826 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Gossypium hirsutum (cotton)
: IMMEDIATE SOURCE:
: CLONE: pMDC-15 (NRRL B-21594)
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 31..1647
: OTHER INFORMATION: /product= "Cotton protox-1 coding
: region"
: OTHER INFORMATION: "region"
: US-08-808-931-15
```

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

alignment_block:

US-09-508-418-2 x US-08-808-931-15 ..

Align seg 1/1 to: US-08-808-931-15 from: 1 to: 1826

```
47  Trrprgrhrrgyserrvalalalyasprtyrthvalprosereral 63
    TTTCAAGCTCCGATCTCCCTCGCCGAGGCTCCACGATTCCTCATATA 179
130  TTTCAAGCTCCGATCTCCCTCGCCGAGGCTCCACGATTCCTCATATA 179
    aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
180  AATGACGGGGGAGAAATCATCATCGGAGATTGGTCATCGTTGGAGGTG 229
    80  LylIeserGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
    GATACAGGAGCTTGTGATTCATTCACGCTCGCCACCAAGCACCCTGAC 279
230  GATACAGGAGCTTGTGATTCATTCACGCTCGCCACCAAGCACCCTGAC 279
    95  .... ProAsnLeuMetValThrGluAlaAlaArgAspArgAlaGlyAs 109
    GTCCGCTTCCAAATGTGATTTGTGACGAGAGCCAGAACCCGTGTGGTCA 329
280  GTCCGCTTCCAAATGTGATTTGTGACGAGAGCCAGAACCCGTGTGGTCA 329
    109  nIleThrThrValGluArgAspGlyTyrLeuTyrPgluGluGlyProAsn 126
    CATCACTACCGCTTGAGAGAGATGATATCTGTGGGAAGAACGCCCAACA 379
330  CATCACTACCGCTTGAGAGAGATGATATCTGTGGGAAGAACGCCCAACA 379
    126  erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
    GTTTTCACCCCTCCGATCTATCTTACACCATGCGCGTGATAGTGATTTG 429
380  GTTTTCACCCCTCCGATCTATCTTACACCATGCGCGTGATAGTGATTTG 429
    143  LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
    AAGGACGATTTGGTTTGTGTGACCCCTAATGACCGCGATTTGTGATATG 479
430  AAGGACGATTTGGTTTGTGTGACCCCTAATGACCGCGATTTGTGATATG 479
    159  pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
    GGAGGGAAGAACTAAGGCTGTGCTCCCAACCAACGACTTGGCCGTTT 529
480  GGAGGGAAGAACTAAGGCTGTGCTCCCAACCAACGACTTGGCCGTTT 529
    176  heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
    TTTGATTTGATGACATTCCTGCGAAACCTTAGGGCTGGGGCTCTATT 579
530  TTTGATTTGATGACATTCCTGCGAAACCTTAGGGCTGGGGCTCTATT 579
    193  GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
    GGCATTTCGGGCTCCCTCCGCTCGGTTATGAAGATCGGTGAGAGATTGT 629
580  GGCATTTCGGGCTCCCTCCGCTCGGTTATGAAGATCGGTGAGAGATTGT 629
    209  LArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
    GCGCGCTAATCTTGTGTGCTGAGGTTTGAACGCTTATGTACCATTTT 679
630  GCGCGCTAATCTTGTGTGCTGAGGTTTGAACGCTTATGTACCATTTT 679
    226  ySerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
    GTTCAGGTCTTATGACAGGGGATCCTTCAAAATTAAAGCATGAACACCA 729
680  GTTCAGGTCTTATGACAGGGGATCCTTCAAAATTAAAGCATGAACACCA 729
    243  PheGlyLysValTyrPlysLeuGluGluThrGlyGlySerIleIleGly 259
    TTTGGAGAGGATGAGAGCTAGAACGATTTGGTGACACATCATTTGTGG 779
730  TTTGGAGAGGATGAGAGCTAGAACGATTTGGTGACACATCATTTGTGG 779
    259  yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
    CACTTTCACGACAAATCCAGAGAGAAATTAAGACACTTAAGCCACCCAG 829
780  CACTTTCACGACAAATCCAGAGAGAAATTAAGACACTTAAGCCACCCAG 829
    276  sProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys 292
    ACCCGGCTGTGCCAAACCGAAGGGCCAAACGTTGGATCTTTGAAG 879
830  ACCCGGCTGTGCCAAACCGAAGGGCCAAACGTTGGATCTTTGAAG 879
    293  GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
```

```
880  GGACTTACCATGCTGCTGAGGCAATGTCTAACAGTTTGGTAGCAATGT 929
309  ulysLeuSerTrrPlysLeuSerSerIleThrLysSerGluLysGlyTyr 326
    AAAATATCTTGTGAAGCTTCCAGATTAACAAATTGGGCAATGGAGGCT 979
930  AAAATATCTTGTGAAGCTTCCAGATTAACAAATTGGGCAATGGAGGCT 979
    326  yHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
    ATACTTGACATTTGAACACCTCGAAGCAATGTATCTTCAGAGTGA 1029
980  ATACTTGACATTTGAACACCTCGAAGCAATGTATCTTCAGAGTGA 1029
    343  SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
    AGTGTGTAATGACATTCATCCATGCTGTCAGTACATCTTGTTCATCC 1079
1030  AGTGTGTAATGACATTCATCCATGCTGTCAGTACATCTTGTTCATCC 1079
    359  OleuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProPov 376
    TCTCTCGGTGCTGCTGAGATGATATCCAAATTTATTTATCTCCAG 1129
1080  TCTCTCGGTGCTGCTGAGATGATATCCAAATTTATTTATCTCCAG 1129
    376  aLGlValAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArg 392
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
1130  TTGCATCAGTCAACAGTCTCTATCCAAAGAACCATTCGAAAGAAATGT 1179
    393  LeuValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrG 409
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
409  nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsn 426
    AGGAATTTGAACCTTTAGGACGATATACAGTTTCAATCATTCCCAATC 1279
1230  AGGAATTTGAACCTTTAGGACGATATACAGTTTCAATCATTCCCAATC 1279
    426  rGAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyLysAlaLys 442
    GAGCTCATCTGCGACGAGGTGTCTCTGAACATACATAGAGAGAGCTACC 1329
1280  GAGCTCATCTGCGACGAGGTGTCTCTGAACATACATAGAGAGAGCTACC 1329
    443  AspProGluIleLeuSerLysThrGluSerGlnLeuValGluValAlas 459
    AACCTGGAAATTTTTCCAAGACTGAAGGGAAGCTTGTGAAGCACTTGA 1379
1330  AACCTGGAAATTTTTCCAAGACTGAAGGGAAGCTTGTGAAGCACTTGA 1379
    459  PARAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
    TCCTGATTTGAGAAATGCTTATTAATCCTAATGCAAGAGATCCTCTTG 1429
1380  TCCTGATTTGAGAAATGCTTATTAATCCTAATGCAAGAGATCCTCTTG 1429
    476  aValAlaGlyAlaArgValTrrProGlnAlaIleProGlnPheLeuValG 492
    TTTTGGGTGTGAAGATATGCCCCAAAGCCATTCCACAGTTGTTGGTGT 1479
1430  TTTTGGGTGTGAAGATATGCCCCAAAGCCATTCCACAGTTGTTGGTGT 1479
    493  HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLe 509
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
1480  CATTTGGATCTCCTTGTATGATGCAAAATGCTCTCAGGAGATTCTGGGTT 1529
    509  uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
1530  TCATGAGACTTCTTCTGGGGGCAACTATGTATCTGTGTGGGCTTGTGGAC 1579
    526  rGcysValGluGlyAlaTyrGluValAlaSerGlyValAlaThrGlyPheLeu 542
    GGTGTGTGGAGGCTGTACGAGGTTGACGTGAAGTGAAGCAATTCCTG 1629
1580  GGTGTGTGGAGGCTGTACGAGGTTGACGTGAAGTGAAGCAATTCCTG 1629
    543  SerArgTyrAlaTyrLys 548
    TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT  TTTT
1630  TCACAAATATGCAATCAAA 1647
```

seq_name: /cgn2_6/plodata/1/lna/6A_COMB.seq:US-08-808-323-15

seq_documentation_block:

; Sequence 15, Application US/08808323

; Patent No. 6018105

; GENERAL INFORMATION:

; APPLICANT: Johnson, Marie

; APPLICANT: Volrath, Sandra

; APPLICANT: Ward, Eric

; TITLE OF INVENTION: Promoters from Plant

; TITLE OF INVENTION: Protoporphyrinogen Oxidase Genes


```

1  NUMBER OF SEQUENCES: 26
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: NO. 6018105arts Corporation
4  STREET: 520 White Plains Road, P.O. Box 2005
5  CITY: Tarrytown
6  STATE: NY
7  COUNTRY: USA
8  ZIP: 10591-9005
9  COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: PatentIn Release #1.0, Version #1.30
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/808,323
16 FILING DATE:
17 CLASSIFICATION: 800
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 60/012,705
20 FILING DATE: 28-FEB-1996
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER: US 60/013,612
23 FILING DATE: 28-FEB-1996
24 PRIOR APPLICATION DATA:
25 APPLICATION NUMBER: US 60/020,003
26 FILING DATE: 21-JUN-1996
27 ATTORNEY/AGENT INFORMATION:
28 NAME: Meligs, J. Timothy
29 REGISTRATION NUMBER: 38,241
30 REFERENCE/DOCKET NUMBER: CGC 1846
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (919) 541-8587
33 TELEFAX: (919) 541-8689
34 INFORMATION FOR SEQ ID NO: 15:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 1826 base pairs
37 TYPE: nucleic acid
38 STRANDEDNESS: single
39 TOPOLOGY: linear
40 MOLECULE TYPE: cDNA
41 HYPOTHEetical: NO
42 ANTI-SENSE: NO
43 ORIGINAL SOURCE:
44 ORGANISM: Gossypium hirsutum (cotton)
45 IMMEDIATE SOURCE:
46 CLONE: pMDC-15 (NRRL B-21594)
47 FEATURE:
48 NAME/KEY: misc_feature
49 LOCATION: 31..1647
50 OTHER INFORMATION: /product= "cotton protox-1 coding
51 OTHER INFORMATION: region"
52 US-08-808-323-15
53
54 Alignment_scores:
55 Quality: 2113.00 Length: 506
56 Ratio: 4.477 Gaps: 1
57 Percent Similarity: 93.281 Percent Identity: 79.249
58
59 Alignment_block:
60 US-09-508-418-2 x US-08-808-323-15 ..
61
62 Align seg 1/1 to: US-08-808-323-15 from: 1 to: 1826
63
64 47 TrrpgrthargcysservAlAlaAlaAspTyrThryValProserSerAl 63
65 :::::::::::::::::::::::::::: |||::: |||||
66 130 TTCAAGCTCCGATCTCCCTCGCGAGGmCCACGCTTCCTCATCTAA 179
67 :::::::::::::::::::::::::::: |||::: |||||
68 63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
69 ::::::::::| | | | | :::::: | | | | | | | | | | | | | |
70 180 AATGACAGGGGAGAAATCATCATCCATCGCGGATTCGCTATCGTTGAGAGTC 229
71
72 80 LylleSerGlyLeuGlyCylleAlaGlnValMetSerAlaAsnTyr..... 94

```

| | | | |
|-----|---|------|--|
| 230 | GTATCAGTGGACCTTGCATTTGGTCAAGGCTCTCGCCACCAACGACCTGCAC | 279 | |
| 95 | ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys | 109 | |
| 280 | GTGCGTTCCAAATGTATATTGTACGGAGGCCAGAACCCGTGTGGTGGCAA | 329 | |
| 109 | nileThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsn | 126 | |
| 330 | CATCACTACCGCTTGAGAGNATGATATCTGTGGGAAGAGGCCCAACA | 379 | |
| 126 | erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu | 142 | |
| 380 | GTATTACGCCCTCCGATCCTATTCTCAACCAATGGCCGGATATGATGATTG | 429 | |
| 143 | LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr | 159 | |
| 430 | AAAGACGATTGGTTTATGGAGACCCTAAAGCAACCCGAGATTGTACTATG | 479 | |
| 159 | plysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP | 176 | |
| 480 | GGAGGSAAAACTAAGAGCCTGTGCCCTCCAAAGCCACACCGACTTCGCTTT | 529 | |
| 176 | heAspLeuMetSerLleProGlyLysLeuArgAlaGlyPheGlyProIle | 192 | |
| 530 | TTGATTGTATGATGACATTCCTGTGAACAACTTATGAGCGTGGTCCGGCTATT | 579 | |
| 193 | GlyLeuArgProSerProProGlyHisGluGlnSerValGluGlnPheVal | 209 | |
| 580 | GGCATTCGGCCCTCCCCCTCCGGGTTATGAAGAAATCGGTGAGAGATTGTT | 629 | |
| 209 | lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC | 226 | |
| 630 | GGCGCGTATCTTGCTGCTGAGGATTTTGAACCTTATTTGAACCACTTTT | 679 | |
| 226 | ysSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla | 242 | |
| 680 | GTTCAGGTTTATGACAGGGGATCTTTCAAAATTAAGCATGAAGACAGA | 729 | |
| 243 | PheGlyLysValTyrPlyLysGluGluThrGlyGlySerLleGlyGly | 259 | |
| 730 | TTTGGAGAGTATGGAAGCTGAAGAGATTTGGTGGCAGCATTTGGTGG | 779 | |
| 259 | yrnhrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg | 276 | |
| 780 | CACCTTCAAGACATTCAGAGGAAATTAAGACACTTAAGCCACCCACAG | 829 | |
| 276 | spProArgLeuProLysProLysGlyGlnThrValGlySerPheArgLys | 292 | |
| 830 | ACCCGCGTCTGCCAAACCCAGAGGCCCAAAACACTTGGATCTTTTGGAG | 879 | |
| 293 | GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe | 309 | |
| 309 | uLysLeuSerTrpLysLeuSerSerLleThrLysSerGlyuysGlyGlyT | 326 | |
| 930 | AAATATTACTTGGAGAGCTTCCAGTATTTACCAATTTGGCAATGGAAGGT | 979 | |
| 326 | yrnLysLeuThrTyrGluThrProGluGlyValAlaValSerLeuInsery | 342 | |
| 980 | ATAACTTGACATTTGAAGAACCTCGAAGAGATGGTATCTTCCAGAGTGA | 1029 | |
| 343 | SerLleValMetThrValProSerTyrValAlaSerAsnLleLeuArgPr | 359 | |
| 030 | AGTGTGTTAAATGACCAATTCATCCCATGTTGGCAGTAAGTCTTGCCATCC | 1079 | |
| 359 | OleuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProP | 376 | |
| 080 | TCTCGCGGCTGCTCTCGACATGATTCATTCACAAATTTATATCTCTCAG | 1129 | |
| 376 | alGlyAlaValAlaThrLleSerTyrProGluGluAlaIleArgAspGluArg | 392 | |

```

1130 TTGCATCAGTCACAGTCTCTCTATCCAAAGAACCATTCGAAAGATGT 1179
393 LeuValAspGlyLeuLeuGlyPheGlyGlnLeuHisProArgThrG1 409
1180 TTGATTGATGTGTAAGTAAAGGGTTGGCCAGTTCACCCAGCAGCA 1229
409 nglyValGlnThrLeuGlyThrIleTyrSerSerLeuPheProAsnA 426
1230 AGGAATTGAAACTTAAAGGAGATATACACTTCATCCTTCCCAATC 1279
426 gAlaProGlyGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
1280 GAGTTCATCTGCGAGGGTGTGTCTTGTGAACATACATAGAGAGGTACC 1329
443 AsnProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAla 459
1330 AACACTGGAATTTTGTCCAAAGACTGAAGGGGAACTGTAGAACAGTTGA 1379
459 PArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476
1380 TCGGATTGTGAAAGAAATGCTTATTAATCCTATTCAGAAAGATCCTTG 1429
476 aValGlyValArgValTyrProGlnAlaIleProGlnPheLeuValGly 492
1430 TTTGGGTGATAGAGTATAGGCCAAAGCCATTCACAGTTCCTGTGTGT 1479
493 HisLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGly 509
1480 CATTTGGCTCTCCTTGATAGTCAAAAATGGCTCTCAGGATTCCTGGGTT 1529
509 uGluGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGly 526
1530 TCATGAGACTGTTCTTGGGGCAACTATGATCTGCTGGCATTAGAC 1579
526 rGcysValGluGlyAlaTyrGlnValAlaSerGluValThrGlyPheLeu 542
1580 GGTGTGTGAAGTGTCTACGAGGTCGACGCTGAAGTGAAGAAATTCCTG 1629
543 SerArgTyrAlaTyrLys 548
1630 TCACAAATATGATACAA 1647
seq_name: /cgn2_6/prodata/1/lna/6A_COMB.seq:US-09-050-603A-15
seq_documentation_block:
; Sequence 15, Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6023012artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,603A
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Mels, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1826 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gossypium hirsutum (cotton)
; IMMEDIATE SOURCE:
; CLONE: pMDC-15 (NRRL B-21594)
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 31..1647
; OTHER INFORMATION: /product= "Cotton protox-1 coding
; OTHER INFORMATION: sequence"
; US-09-050-603A-15

```

```

alignment_scores:
Quality: 2113.00 Length: 506
Ratio: 4.477 Gaps: 1
Percent Similarity: 93.281 Percent Identity: 79.249

```

```

alignment_block:
US-09-508-418-2 x US-09-050-603A-15 ..
Align seg 1/1 to: US-09-050-603A-15 from: 1 to: 1826

```

```

47 TTPArgThrArgCysSerValAlaLysAspTyrThrValProSerSerAl 63
130 TTCAGCTCCGATGCTCCCGCGAGGGTCCACGATTCCTCATCTAA 179
63 aValAspGlyGlyProAlaAlaGlnLeuAspCysValIleValGlyAla 80
180 AATCGAGGGGGAGAAATCATCCATCCGCGAATTCGCTATCGTGGAGTGC 229
80 LylLeSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
230 GTATCAGTGCATTTTGCATTTGCTCAAGCTCTGCCACCAAGCACCGTGAC 279
95 .....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyLys 109
280 GTGCTTCCAATGTGATTTGACGAGGCGACAGACCGTGTGCTGCAG 329
109 nIleThrThrValGluArgAspGlyTyrLeuTyrGluGluGlyProAsn 126
330 CATCATTACCGCTTGAGAGAGATGATATCTGTGGAAAGGCCCCAACA 379
126 erPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
380 GTTTTCAGCCCTCCGATCTATTTAACAATGCGCGTGGATAGTATG 429

```

```

1330 ANACACGGAAATTTTGTCCAAAGCTGGAAGGGGAATTTGTAGACAGATTGA 1379
4455 pArgAspLeuArgLysMetLeuIleTyrSProlYSAlaGlnAspProLeuY 476
1380 TCGTGTATTTGAGAAAATGCTTTAAATCTTAATGCAAAAGGATCTCTTG 1429
476 aAlaGlyValArgValTrrProGlnAlaIleProGlnPheLeuValGly 492
1430 TTTTGGGTGTAAAGTATGTGGCCAAACCCATCCACAGTTCTTGTTGGT 1479
493 HSLaenAsPThrIeusSerThrAlaLysAlaAlaMetAsnSPaSnGlyLe 509
1480 CATTGGATGATCCTTGATAGTGCAGAAAATGGCTCTCAGGGATTTGGGTT 1529
509 uGluGlyLeuPheLeuGlyGlyLysnTyrValISerGlyValAlaLeuGlyA 526
1530 TCAATGACGTCTTCTTGGGGGCAACTATGATCTGATGCGTGGCATTAAGC 1579
526 rGcYSaIGluGlyAlaTyrGlyValAlaIaSerGluValAlaThrGlyPheLeu 542
1580 GCGTGTTGGAGAGGTGCTTAGAGAGTTGCACACTGAAGTGAAGAGATTTCTG 1629
543 SerArgTyrAlaTyrLys 548
1630 TCACAAATATGCATACAA 1647

seq_name: /cgn2_6/ptodata/1/ina/6A_COMB.seq:US-09-102-420B-15

seq_documentation_block:
Sequence 15, Application US/09102420B
Patent No. 6084155
GENERAL INFORMATION:
APPLICANT: Volirath, Sandra
APPLICANT: Johnson, Marie
APPLICANT: Ward, Eric
APPLICANT: Helfetz, Peter
TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6084155artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/102,420B
FILING DATE: 22-JUN-1998
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/059,164
FILING DATE: 13-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/050,603
FILING DATE: 30-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/126,430
FILING DATE: 11-MAR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/808,931
FILING DATE: 28-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612

```

```

FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/472,028
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ. ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1826 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEetical: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Gossypium hirsutum (cotton)
IMMEDIATE SOURCE:
CLONE: pMDC-15 (NRRL B-21594)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..1647
OTHER INFORMATION: /product= "Cotton protox-1 coding
: OTHER INFORMATION: sequence"
US-09-102-420B-15

```

```

alignment_scores:
Quality: 2113.00      Length: 506
Ratio: 4.477          Gaps: 1
Percent Similarity: 93.281      Percent Identity: 79.249

```

alignment_block:

US-09-508-418-2 x US-09-102-420B-15 ..

```

Align seg 1/1 to: US-09-102-420B-15 from: 1 to: 1826

47 TTPAAGTATGCGSeraValAlaLysAspTyrThrValProSerSerAl 63
   ::::: |||||||:::||||: ||::: |||||
130 TTCAAGCTCCGATGCTCCCTCCGCGAGGGTCCACGATTTCCTCATCTAA 179
   ::::: |||||||:::||||: ||::: |||||
63 aValAspGlyGlyProAlaAlaGluLeuAspCysValIleValGlyAlaG 80
   ::::: |||||||:::||||: ||::: |||||
180 ATTCACGCGGAGAGATCATCATCGCGATTGCGTCATCGTTGAGAGTG 229
   ::::: |||||||:::||||: ||::: |||||
80 yLisSerCylLeuCysIleAlaGluValMetSerAlaAsnTyr..... 94
   |||||||:::||||: ||::: |||||
230 GTATCAGTGGACTTGCATTCCTCAAGCTTCGCCACCAACACGCGTGAC 279
   ||::: |||||
95 .....ProAsnLeuMetValThrLysAlaArgAspArgAlaGlyGlyAs 109
   ||::: |||||
280 GTGCGTCCCAATGATGATGTGACGAGGCCAGAGACCGTGTGTGGCAA 329
   ||::: |||||
109 nIleThrThrValGluArgAspGlyTyrLeuThrPgluGluGlyProAsn 126
   |||||||:::||||: ||::: |||||
330 CATCACTACCGCTTGAGAGAGATGATATCTGTGGAGAGAGAGCCCAACA 379
   |||||||:::||||: ||::: |||||
126 ePheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeu 142
   |||||||:::||||: ||::: |||||
380 GTTTCACGCCCTCGATCTATTATTAACCATGCGCGTGATAGTAGTTG 429
   |||||||:::||||: ||::: |||||
143 LysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTr 159
   |||||||:::||||: ||::: |||||
430 AAGGACGATTGGTTTGTAGTGACCTAATGACACCGCATTTGTACTATG 479

```

```

159 pLysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPheP 176
   ||::: |||||
480 GAGGAGAAACCTAAGGCGCTGTGCCCTCCAAGCCAAACGACTTGGCGCTTT 529
   ||::: |||||
176 heAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 192
   |||||||:::||||: ||::: |||||
530 TTGATTTGATGACCATTTGCTGAGAAACCTAAGGCTTGGTTCGGGGCTATT 579
   ||::: |||||
193 GlyLeuArgProSerProProGlyHisGluGluSerValGluGlnPheVa 209
   ||::: |||||
580 GGCAATCGGCCCTCCCTCCGCGGTATGAAGAAATCGGTGGAGAGATTGTT 629
   ||::: |||||
209 lArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheC 226
   |||||||:::||||: ||::: |||||
630 GCGCCGTAATCTTGCTGCTGAGGTTTGTGAACGCTTATTGAACCATTTT 679
   ||::: |||||
226 ySserGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAla 242
   |||||||:::||||: ||::: |||||
680 GTTCAGGTGTTTATGCAAGGGATCCTTCAAAATTAAGCATGAAGCAGCA 729
   ||::: |||||
243 PheGlyLysValThrPlyLysLeuGluGluThrGlySerIleIleGlyGl 259
   |||||||:::||||: ||::: |||||
730 TTTGGAAGAGTATGAGAGCTAGAGAGATTGTTGCGACATCATTTGTTGG 779
   ||::: |||||
259 yThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgA 276
   |||||||:::||||: ||::: |||||
780 CACTTTCACACATCCAGAGAGAGAAATTAAGACCTTAAGCCACCCAGAG 829
   ||::: |||||
276 sProArgLeuProLysProLysGlyGlnThrValLysSerPheArgLys 292
   |||||||:::||||: ||::: |||||
830 ACCCGCGCTCGCCAAACCCAAAGGCCAAACAGTTGATCTTTTAAAGAAG 879
   ||::: |||||
293 GlyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLe 309
   |||||||:::||||: ||::: |||||
880 GGACTTACCATGCTGCTGAGGCAATTCCTACACATTTGGGTGACATGT 929
   ||::: |||||
309 uLysLeuSerThrLysLeuSerSerIleThrLysSerGluLysGlyLys 326
   |||||||:::||||: ||::: |||||
930 AAATTTATCTTGGAGGTTTCCAGTATTACCAAAATTTGGCAATGAGAGGT 979
   ||::: |||||
326 yHisIleThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 342
   |||||||:::||||: ||::: |||||
980 ATAACTTGACATTTGAAACACCTGAAGGAATGATCTCTTCAGATTAAGA 1029
   ||::: |||||
343 SerIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPr 359
   ||::: |||||
1030 AGCTTTATTAAGACATTCATCCATGTTGCCAGATTAATCTGTTGATGC 1079
   ||::: |||||
359 oLeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProV 376
   |||||||:::||||: ||::: |||||
1080 TCTCTCGCGTGTGTCGACATGCATTAATCCCAATTTTATTAATCCTCCAG 1129
   ||::: |||||
376 aLgIlyAlaValThrIleSerTyrProGluGluAlaIleArgAspGluArg 392
   |||||||:::||||: ||::: |||||
1130 TTGCATGACGACAGCTCTCTATCCAAAGAACCACTTGGAAAGAAATGT 1179
   ||::: |||||
393 LeuValAspGlyLysLeuLysGlyPheGlyGlnLeuHisProArgThrGl 409
   ||::: |||||
1180 TTGATTTATGTCGACTTAAGGGGTTTGGCCAGTTGCACCCAGCGAGCCA 1229
   ||::: |||||
409 nGlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnA 426
   |||||||:::||||: ||::: |||||
1230 AGGAATGGAACCTTAGGAGCATATACAGTATCATCACTTTTCCCAATC 1279
   ||::: |||||
426 rGluArgProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 442
   |||||||:::||||: ||::: |||||
1280 GAGCTTCATCTGCGAGGGTGTCTCTTGAATCAATAGAGAGAGGTACCC 1329
   ||::: |||||
443 AsnProGluIleLeuSerLysThrGluSerSerLeuValGluValAlaLys 459
   |||||||:::||||: ||::: |||||
1330 AACACGTGGAATTTTGTCCAAAGACTGAAGGGAGACCTTGTAAGACATTGA 1379
   ||::: |||||
459 pArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuV 476

```

```

|||||
1380 TCCTATTGAGAAAAATGCTTATTAATCTTAATGCAAGGATCTCTTG 1429
476 aIvalGlyValArgValIrrProGlnAlaIleProGlnPheLeuValGly 492
|||||
1430 TTTGGGTGTAAGAGTATGGCCAAAGCCATTCCACAGTTCTTGTTGCT 1479
493 HistLeuSprHLeuSerThrAlaLysAlaAlaMetLsnAspAsnGlyLe 509
|||||
1480 CATTGGATCTCTTGGATAGTCAAAAAATGGCTCTCAAGGATTCCTGGGTT 1529
509 uGIuGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyA 526
|||||
1530 TCATGGAGCTGTTCTTGCGGCACTATGTAATCTGGTGCTGATTAAGAC 1579
526 rGcysValGluGlyAlaTyrGluValAlaSerGluValIrrHcGlyPheLeu 542
|||||
1580 GGTGTGGAGAGTGTCTACGAGGTCTGACGTGAAGTGAAGAAATTCCTG 1629
543 SerArgTyrAlaTyrLys 548
|||||
1630 TCACAAATATGCATACAA 1647

```

seq_name: /cgn2_6/plodata/1/lna/5A_COMB.seq:us-08-472-028A-1

seq_documentation_block:

```

; Sequence 1, Application US/08472028A
; Patent No. 5767373
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Volrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; TITLE OF INVENTION: Oxidase Enzyme Activity In Eukaryotic Organisms
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472.028A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: CGC 1748/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8614
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1719 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1644
; OTHER INFORMATION: /note="Arabidopsis protox-1 cDNA;
; OTHER INFORMATION: sequence from pMDC-2"
; US-08-472-028A-1

```

alignment_scores:
 Quality: 2077.00 Length: 539
 Ratio: 4.230 Gaps: 4
 Percent Similarity: 91.095 Percent Identity: 75.139

alignment_block:
 US-09-508-418-2 x US-08-472-028A-1 ..

Align seg 1/1 to: US-08-472-028A-1 from: 1 to: 1719

```

19 SerSerProLeuAlaPheLeuLsnArgThrSerPheIleProPheSerSe 35
|||||
28 TCATGAGATTTATCTCTCTCCGTCGACGACTCAATCCCTTCTTCGCGC 77
35 rIleSerLysArgAsnSerValAsnGlyLysnGlyTrp.....ArgT 49
|||||
78 GTTTTCGAAGCCCAT...CTCGATTAAATGTTATTAAGCCTTAGAC 124
49 hrArgCysSerValAlaLysAspTyrThrValProSerSerAlaValasp 65
|||||
125 TCAGTGTTCAGTGGCGCGTGGACCAACCGCTGATCTCAAAAAATCCGA 174
66 GlyGlyProAlaAlaGluLeu.....AspCysValIleValGlyAlaG 80
|||||
175 GCGGAGAGAGGCCACCAACCATCAACGAGATGTGTGATGTGCGGAGAG 224
80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||||
225 TATTAGTGTCTTTGATGCTGCTCAGGCGCTTCTACTAAGCAATCCTGATG 274
95 ....ProAsnLeuMetValIrrHcGluAlaArgAspArgAlaGlyLysn 109
|||||
275 CAGCTCCGAATTAATTATGACCGAGCTAAGATGATGTGTGAGGCAAC 324
110 IleThrThrValGluArgAspGlyTyrLeuTrpGluGluGlyProAsnSe 126
|||||
325 ATTATCACTCGTGAAGAGATGTTTCTCTGGAAGAGGCCCAATAG 374
126 rPheGlnProSerAspProMetLeuThrMetAlaValAspCysGlyLeuL 143
|||||
375 TTTTCAACCGTGTGATTCCTATGCTCATAGTGTGATAGATAGTGGTTTGA 424
143 ysAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeuTrp 159
|||||
425 AGCATGATTTGGTGTGGAGATCCTACTGCGCAAGTTGTGTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhePh 176
|||||
475 AATGGAAATTAAGCCGCTTCATCAAGAGCTAACAGACTTACCGTTCTT 524
176 easPleuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
|||||
525 TGATTGATGATATTGTTGGGGAATATAGACCTGTTTGGTGACACTTG 574
193 lYleuArgProSerProGlyHisGluGluSerValGluGlnPheVal 209
|||||
575 GCATTGACCGCTCACCTCCAGGTCGTGAAGAAATCTGTGAGAGAGTTGTA 624
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
|||||
625 CGGCGTAACCTCGGTGATGAGAGTCTTGAAGCCCTGATGAACCGTTTGG 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaIleP 243
|||||
675 TTCAGGTGTTATGCTGTGATCTCTCAAACTGACATGAAGACAGCGCT 724
243 heGlyLysValIrrPylsLeuGluGluTrpGlyGlySerIleIleGlyGly 259
|||||
725 TTGGGAAGTTTGGAACTAGACAAATGCTGGAAGCATATATAGTGTGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArgAs 276
|||||

```

```

775 ACTTTAAGCAATTCAGAGAGGAAAAAGCTCCAGAGCAGACAGAGA 824
276 PProArgLeuProLysProLysGlyInThrValGlySerPheArgLysG 293
825 CCCCGCGCTGCAGAAACACAGAGGCCAAACAGTTGGTTCTTCAGAGAG 874
293 LLeuArgMetLeuProAspAlaLeuSerAlaArgLeuGlySerLysLeu 309
875 GACTTCGAAGTTGGCAGAACCAATATCTCAGAAATTAGGTAGCAAGTT 924
310 LysLeuSerThrLysLeuSerSerLeuThrLysSerGluLysGlyLys 326
925 AAGTTGCTTCGGAAGCTCTCAGATACACRAAGTGGAGAGCGGAGATA 974
326 rHisLeuThrTyrGluThrProGluGlyValSerLeuGlnSerArgS 343
975 CAACCTTAACATATGAGACTCCAGATGCTTATGTTCCGTCAGAGCAAAA 1024
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
1025 GTGTTGAATGACGCGTCCATCTCATGTTCAGAGTGGTCTCTGGCGCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProProVa 376
1075 CTTTCTGAATCTGCTGCAAAATGCACTCTCAAACTATATTATACCAAGT 1124
376 LglValAlaValThrIleSerTyrProGluGluAlaIleArgAspGluArgL 393
1125 TGCAGACAGATATCTATCTGTCACCGAAGAGCAATCCGAAACAGAAATGTT 1174
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
1175 TGATAGATGCGCAACTAAGAGGTTTGGGCAATTGCATCCAGCAGCAAA 1224
410 GlValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
1225 GGAGTTGAACATTAGCAACTATCTACAGCTCTCCACTTTCCAAAATCG 1274
426 gAlaProLysGlyArgValIleLeuLeuAsnTyrIleGlyValAlaLysA 443
1275 CGCACCGCCGCGAAGAAATTTGCTGTGAACACTCAATGGCGGCTTCMA 1324
443 snProGluIleLeuSerLysThrGlnSerGlnLeuValGluValAlaAsp 459
1325 ACACCGGAATTCCTGCAAGCTGAGAGTGAGTGTAGTGAGAGCAGTTGAC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
1375 AGAGATTTGAGGAAATGCTAATTAAAGCTTAATTCGACGATCCACTTAA 1424
476 lValGlyValArgValIleThrProGlnAlaIleProGlnPheLeuValGlyH 493
1425 ATTAGAGAGTGTAGGCTGCTCAAGCCATTCCTCAGTTTCTAGTTGGTTC 1474
493 lIleuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTGATATCTCTGACAGCGCTAAATCATCTAAAGCTCTTCGCGCTAC 1524
510 GlGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGlyAr 526
1525 GAAGGCTATTTTGGTGGCAATTACGTCGCTGTGTAGCGCTTAAGGCGG 1574
526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeuS 543
1575 GTGTGTAGAGGCGCATATGAACCCGCAATTGAGGTCAACAACATTCAATGT 1624
543 erArgTyrAlaTyrLys 548
1625 CACGCTACGCTTACAG 1641

```

```

seq_name: /cgn2_6/ptodata/1/lna/5B_COMB.seq.us-08-808-931-1
seq_documentation_block:

```

```

: Sequence 1, Application us/08808931
: Patent No. 593602
:
: GENERAL INFORMATION:
: APPLICANT: Voltrath, Sandra
: APPLICANT: Johnson, Marie
: APPLICANT: Potter, Sharon
: APPLICANT: Ward, Eric
: APPLICANT: Helfetz, Peter
: TITLE OF INVENTION: DNA Molecules Encoding Plant
: TITLE OF INVENTION: Protoporphylinogen Oxidase and Inhibitor-Resistant Mutants
: TITLE OF INVENTION: Theof
: NUMBER OF SEQUENCES: 35
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 593602artis Corporation
: STREET: 520 white plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: us/08/808,931
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1847
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8687
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1719 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Arabidopsis thaliana
: IMMEDIATE SOURCE:
: CLONE: pMDC-2 (NRRL B-21238)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 31..1644
: OTHER INFORMATION: /product= "Arabidopsis protox-1"
:
: US-08-808-931-1

```

```

alignment_scores:
  Quality: 2077.00      Length: 539
  Ratio: 4.230          Gaps: 4
  Percent Similarity: 91.095  Percent Identity: 75.139

```

```

alignment_block:
  US-09-508-418-2 x US-08-808-931-1 ..

```

```

Align seg 1/1 to: US-08-808-931-1 from: 1 to: 1719

```


Thu Jul 5 14:36:12 2001

us-09-508-418-2.rni

Page 10

ADDRESSEE: No. 6018105artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,323
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Melgs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: CGC 1846
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: pMDC-2 (NRRL B-21238)
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1644
OTHER INFORMATION: /product= "Arabidopsis protox-1"

```

Alignment scores:
Quality: 2077.00      Length: 539
Ratio: 4.230          Gaps: 4
Percent Similarity: 91.095      Percent Identity: 75.139

alignment_block:
US-09-508-418-2 x US-08-808-323-1 ..

Align seg 1/1 to: US-08-808-323-1 from: 1 to: 1719

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSer 35
   |||:::||||::: ||::: |||
28 TCCATGAGATTTACTTCTTCCTCCGCGAGCATCATCGCTTCTCCGTC 77
   |||SerLysArgGlnSerValAsnCysAsnGlyTrp.....ArgT 49
   |||:::|||| ||| ::::: |||
78 GTTTTCGAAGCCCAAT..CTCCGATTAATATTATATAGCCCTTTAGAC 124
   |||ThrGlySerValAlaIalysaPTrpThnValProSerSerAlaValasp 65
   |||:::||||::: |||::: |||::: |||::: |||::: |||:::
125 TCCGTTGTTCAAGTGGCGGGGTGAGACCAACCGCTGCATCTCAAAAATGCA 174

```

[illegible]


```

||||| .....
1075 CTTTCGAATCTGCTGCAATCTCAATCAATATATTATACCCACAGT 1124
376 1G1yAlaValThrIleSerTyrProGlnIleAlaIleArgSpGluArgL 393
1125 TGCAGCAGATCTATCTGCTACCCGAAGAAGCAATCCGAACAGATGTT 1174
393 euValAspGlyLeuLeuGlyPheGlyGlnLeuHisProArgThrGln 409
1175 TGTATAGATGTGTAACATAAGGGTTTGGGCAATGTGCATCCAGCAGCAA 1224
410 G1yValGluThrLeuGlyThrIleTyrSerSerLeuPheProAsnArg 426
1225 GGAGTTGAACTTGTGAACTATGACAGCTCTCACTCTTCCAAATCG 1274
426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyIleAlaLysA 443
1275 CGCACCAGCCCGGAAGAAATTTGCTTTGAACATTCATTCGGGGCTACAA 1324
443 snProGlnIleLeuSerTyrThrGluSerGlnLeuValGluValAsp 459
1325 ACACCGGAATCTGTCCAACTGTGAAAGTGAGTGTGAGAGCAGTTGAC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
1375 AGAGATTGTAGGAAATATCTTAATTAGCTTAATTCGACCGACACTTAA 1424
476 1vaG1yValArgValTyrProGlnAlaIleProGlnPheLeuValGlyH 493
1425 ATTAGAGATGAGGTATGGCTCAAGCATTCCTCAGTTCTTACTGTGTC 1474
493 1sLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGlyLeu 509
1475 ACTTGTATATCTTGACACCGGCTAAATCATCTCAACGCTCTTCGGGCTAC 1524
510 GluGlyLeuPheLeuGlyLysAsnTyrValSerGlyValAlaLeuGlyArg 526
1525 GAAGGGCATTTTGGGTGGCAATTCAGTCCTGTGTAGCTTACAGCGC 1574
526 gCysValGluGlyAlaTyrGluValAlaSerGluValThrGlyPheLeu 543
1575 GTGTGTAGAAAGCGCATATGAAACCGGCAATTCAGTCAACAACTCATGT 1624
543 eArgTyrAlaTyrLys 548
1625 CACGCTACGCTTACAG 1641
seq_name: /cgn2_6/ptodata/1/lna/6A_COMB.seq:us-09-050-603A-1
seq_documentation_block:
; Sequence 1, Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protoporphyrinogen Oxidase
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6023012artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,603A
; FILING DATE: 30-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/808,931
; FILING DATE: 28-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1719 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: PMDC-2 (NRRL B-21238)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31..1644
; OTHER INFORMATION: /product= "Arabidopsis protoco-1"
; US-09-050-603A-1

alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

alignment_block:
US-09-508-418-2 x US-09-050-603A-1 ..
Align seg 1/1 to: US-09-050-603A-1 from: 1 to: 1719

19 SerSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSe 35
||| .....
28 TCCATGAGAGTATCTCTCTCCGCGAGACTCAATGCTTCCGCTC 77
35 rIleSerLysArgAsnSerValAsnCysAsnGlyTyr.....ArgT 49
|:| .....
78 GTTTCGAAGCCCAAT...CTCCGATTAATGTTTATAGCCTCTTAAC 124
|||| .....
49 hArgCysSerValAlaLysAspTyrThrValProSerSerAlaValAsp 65
|||| .....
125 TCCGTTGTTCAGTGGCGGTGACCAACCGTCGATCTTCAAAAATCGAA 174
|||| .....
66 GlyGlyProAlaIleGluLeu.....AspCysValIleValGlyAlaG 80
|||| .....
175 GCGGAGAGGAGCCACCATCATCAGCAGGATGTGTGATGTGCGGGAGG 224
|||| .....
80 yLleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyr..... 94
|||| .....
225 TATGTGTGTCTTTCATCGCTCAAGCGCTTGCTAGACATCTGATG 274

```

```

95  ....ProAsnLeuMetValThrGluAlaArgAspArgAlaGlyGlyAsn 109
275 CTGCTCCGAATTTAATTGTGACGAGCGTAAGCATCGTTGGAGGCGAAC 324
110 TLeThrThrValGluArgAspGlyTyrLeuThrProGluGluProAsn 126
325 ATTATTCACCTGCTGAAGAGATGTTCTCTGGGAGAGAGAGTCCCAATAG 374
126 rPheGluProSerAspPromeLeuThrMetAlaValAspCysGlyLeu 143
375 TTTTCAACCGCTGATCTATGCTCACTATGCTGAGTAGATGAGTGTGGA 424
143 ysaAspAspLeuValLeuGlyAspProAsnAlaProArgPheValLeu 159
425 AGCATGATGTTGCTGTGGAGATCTCTACTGCGCCAGAGTTGTTGTGG 474
160 LysGlyLysLeuArgProValProSerLysLeuThrAspLeuProPhe 176
475 AATGGGAATTTGAGCGCGGTTCCATCCAGACTAACAGACTTACGGTCTT 524
176 eAspLeuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIle 193
525 TGATTTGATGAGTATTGTTGGGAGATAGAGCTGTTTGTGTCACCTTG 574
193 LysLeuArgProSerProProGlyHisGluGluSerValGluGluPhe 209
575 GCATTTCAGCCGTCACCTCCAGGTGCTGAGAGATCTGTGGAGAGTTGTA 624
210 ArgArgAsnLeuGlyGlyLysValPheGluArgLeuIleGluProPhe 226
625 CGGCTTAACCTCGGTGATGAGGTTTGTGAGCGCTGATTGAACGGTTTG 674
226 sSerGlyValIleValGlyAspProSerLysLeuSerMetLysAlaAla 243
675 TTCAGGTGTTTATGCTGCTGATCCCTTCAAACTGACATGAAACGAGCG 724
243 heGlyLysValIlePheLysLeuGluGluThrGlyGlySerIleGly 259
725 TTGGGAAGGTTTGGAACTAGAGCAAAATGCTGGAAGCATTAATAGTGT 774
260 ThrPheLysAlaIleLysGluArgSerSerThrProLysAlaProArg 276
775 ACTTTTAAGCAATTCAGAGAGAGAAACGCTCCCAAGCAGAGAGAGA 824
276 ProArgLeuProLysProLysGlyGluThrValGlySerPheArgLys 293
825 CCCGGCGCTGCCAAAACACAGGCGCAACAGTGTCTTTCAGAGAGG 874
293 LysLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLys 309
875 GACTTCGAATGTTGCCAGAGCAAAATCTGCAAGATTAGTGAAGCAAGT 924
310 LysLeuSerTrpLysLeuSerSerIleThrLysSerGlyLysGlyTyr 326
925 AAGTTGTCTTGGAGCTCTCAGGTAATCACTAGCTGAGAGCGAGAGATA 974
326 rHisLeuThrTyrGluThrProGluGlyValAlaSerLeuGlnSerArg 343
975 CAACCTTAACATATGAGACTCAGATGCTTAAGTTTCGCGGACAGCAADA 1024
343 eLysLeuValMetThrValProSerTyrValAlaSerAsnIleLeuArg 359
1025 GTGTGTATATGACGGTGCATCTCATGTTGCAAGTGTCTCTGGCGCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAsnPheTyrTyrProPro 376
1075 CTTTTCGATCTGCTGCAAAATGCACTCTCAAAACTATATTACCCACAGT 1124
376 LysLeuValAlaThrIleSerTyrProGluGlyAlaIleArgAspGly 393
1125 TGCAGCAGATATCTCTGTAACCCGAAAGCAAAATGCAACAGAAATGTT 1174
393 euValAspLysLysLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409

```

```

1175 TGATGATGCTGACTAAAGGTTTGGGCAATTCGATCCAGCGACGAA 1224
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsn 426
1225 GGAGTTGAAACATTAGAGACTATCTACAGCTCTCACTCTTCCAAATGC 1274
426 gaLysProLysGlyArgValLeuLeuLeuAsnTyrIleGlyGlyAlaLys 443
1275 CGACCGCCCGGAAAGATTTGCTGTGAGTACATTCGCGGGGTCTACA 1324
443 snProGluLeuSerLysThrGluSerGlnLeuValGlyValAlaAsp 459
1325 ACACCGGAATTTCTGTCAGAGTCTGAAAGTGAAGTGAAGTGAAGTGC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeu 476
1375 AGAGATTGAGGAAATGCTAATTAGCCTAATTGCAACGATCCACTTAA 1424
476 ValGlyValArgValTrpProGlnAlaIleProGlnPheLeuValGly 493
1425 ATTAGAGATTAGGATGATGCTTCAGCCATTCCTCAGTTTCAAGTTGTC 1474
493 LysLeuAspThrLeuSerThrAlaLysAlaAlaMetAsnAspAsnGly 509
1475 ACTTGATATCTTGACACGCGTAATCATCTTAACGTCCTCGGCTTAC 1524
510 GlnGlyLeuPheLeuGlyGlyAsnTyrValSerGlyValAlaLeuGly 526
1525 GAAGGGCTATTTTGGTGGCAATTAAGTGGCTGGTGGTGGCTTACGGCG 1574
526 gCysValGluGlyAlaTyrGlyValAlaAspGlyValThrGlyPheLeu 543
1575 GTGTGTAGAGGCGCATATGAACCGGATTTGAGTGAACAACTCATGT 1624
543 eArgTyrAlaTyrLys 548
1625 CACGGTACGCTTACAG 1641

seq_name: /gcn2_6/ptodata/1/lna/6A_COMB.seq:US-09-102-420B-1
seq_documentation_block:
; Sequence 1, Application US/09102420B
; Patent No. 6084155
; GENERAL INFORMATION:
; APPLICANT: Voltrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINOGEN
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6084155artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,420B
; FILING DATE: 22-JUN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/059,164
; FILING DATE: 13-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/050,603

```

```

? FILING DATE: 30-MAR-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/126,430
? FILING DATE: 11-MAR-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/808,931
? FILING DATE: 28-FEB-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/472,028
? FILING DATE: 06-JUN-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Meigs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: CGC 1847/CIP3
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
? INFORMATION FOR SEQ ID NO: 1:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1719 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHEetical: NO
? ANTI-SENSE: NO
? ORIGINAL SOURCE:
? ORGANISM: Arabidopsis thaliana
? IMMEDIATE SOURCE:
? CLONE: PMDC-2 (NRRL B-21283)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 31..1644
? OTHER INFORMATION: /product= "Arabidopsis protox-1"
US-09-102-420B-1

alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139

alignment_block:
US-09-508-418-2 x US-09-102-420B-1 ..
Align seq 1/1 to: US-09-102-420B-1 from: 1 to: 1719

19 SerSerProleuAlaPheIleuAsnAlaGlyThrSerPheIleProPheSerSe 35
||| |||:|||||:|||||: |||: |||
28 TCCATGAGATTATCTCTTCCTCCGTCGACGACATCGCTCTTCCTC 77
||| |||:|||||:|||||: |||: |||
35 rIleSerLyArgAsnSerValAsnCyAsnGlyTrp.....ArgT 49
||| |||:|||||:|||||: |||: |||
78 GTTTCGAACCCCAAT...CTCCGATTAAATGTTTATACGCTCTTAGAC 124
||| |||:|||||:|||||: |||: |||
49 hrArgCySerValAlaIleAspTyThrValProSerSerAlaValAsp 65
||| |||:|||||:|||||: |||: |||
125 TCCGTGTTCAGTGGCCGCGGACCAACGCTCGATCTTCAAAAACGAA 174
||| |||:|||||:|||||: |||: |||
66 GlyGlyProAlaIleGluLeu.....AspCyValIleValGlyAlaG 80
||| |||:|||||:|||||: |||: |||
175 GCGGAGGAGGACACACCATCACGACGATGTGTGTCGCGGAGAG 224
||| |||:|||||:|||||: |||: |||
80 yIleSerGlyLeuCyIleAlaGluValMetSerAlaAsnTy.....94
||| |||:|||||:|||||: |||: |||
```

```

|||||:|||||:|||||: |||: |||
225 TATTAGTGTCTTGATCGCTCAGCGCTTCTACTAACCATCCGATG 274
95 ....ProAsnLeuMetValThrGluAlaArgAspAlaGlyAsn 109
|||||:|||||:|||||: |||: |||
275 CTGCTCCGAATTTAAATGTGACCGAGCTAAAGATGCTGTGGAGGCAAC 324
|||||:|||||:|||||: |||: |||
110 IleThrThrAlaGluArgAspGlyTyIleTrpGluGluGlyProAsnSe 126
||| |||:|||||:|||||: |||: |||
325 ATTATCATCTGTGAAGAGATGGTTCCTGGGAGAAAGGTCCCAATG 374
|||||:|||||:|||||: |||: |||
126 rPheGluProSerAspProMetLeuThrMetAlaValAspCyGlyLeu 143
|||||:|||||:|||||: |||: |||
375 TTTTCAACGCTCGATCCATCTACTATGAGGTAGATAGTGTTTA 424
|||||:|||||:|||||: |||: |||
143 yAspAspLeuValIleGlyAspProAsnAlaProArgPheValLeuTrp 159
|||||:|||||:|||||: |||: |||
425 AGATGATTTGGTGTGGAGATCCACTGCGCAAGTTTGTGTGTGG 474
|||||:|||||:|||||: |||: |||
160 LySGlyLyLeuArgProValProSerLyLeuThrAspLeuProPhe 176
|||||:|||||:|||||: |||: |||
475 AATGGGAATTTGAGCCGCTTCATGAAAGCTAACGACTTACCGTTCT 524
|||||:|||||:|||||: |||: |||
525 TGATTTGATGATGATATTGGTGGAAATTAGAGCTGTTTGGTGCAC 574
|||||:|||||:|||||: |||: |||
193 LyLeuArgProSerProProGlyHisGluGluSerValGluGluPheVal 209
|||||:|||||:|||||: |||: |||
575 GCATTTGACCGTCCCTCCAGGTCGTAAGAAATCTGTGAGGAGTTTGA 624
|||||:|||||:|||||: |||: |||
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPhe 226
|||||:|||||:|||||: |||: |||
625 CGGCGTAACCTCGGTATGATGAGTTTGGAGCCGCTGATTAACCGTTTG 674
|||||:|||||:|||||: |||: |||
226 sSerGlyValIlyrValGlyAspProSerLyLeuSerMetLyAlaAla 243
|||||:|||||:|||||: |||: |||
675 TTCAGGTGTTTATGCTGTGATCTTCAAACTGACGATGAAGAAGCAG 724
|||||:|||||:|||||: |||: |||
243 heGlyLyValIlyrLyLeuGluGluThrGlyGlySerIleIleGly 259
|||||:|||||:|||||: |||: |||
725 TTGGAGAGTTTGGAACTAGAGCAAAATGTGGAGCAATTAATAGTGGT 774
|||||:|||||:|||||: |||: |||
260 ThrPheLyAlaIleLySGlyLeuArgSerSerThrProLyAlaProArg 276
|||||:|||||:|||||: |||: |||
775 ACTTTTAAGCAATTCAGAGAGAAACGCTCCCAAGCGAGAACGAGA 824
|||||:|||||:|||||: |||: |||
276 pProArgLeuProLySProLyGlyGluThrValGlySerPheArgLySG 293
|||||:|||||:|||||: |||: |||
825 CCCGCCCTGCCAAAACACAGAGGCCAAACAGTTGTTTCAAGAAAG 874
|||||:|||||:|||||: |||: |||
293 LyLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLyLeu 309
|||||:|||||:|||||: |||: |||
875 GACTTCGATGTTGCCAGAGCAATATCTGCAGATTTGGTAGCAAACTT 924
|||||:|||||:|||||: |||: |||
310 LyLeuSerTrpLyLeuSerSerIleThrLySerGlyLySGlyGly 326
|||||:|||||:|||||: |||: |||
925 AAGTGTCTTGAAGCTCTCAGGTATCACTAAGCTGAGAGGAGGAGTA 974
|||||:|||||:|||||: |||: |||
326 rHisLeuThrTyArgIuThrProGluGlyValValSerLeuGlnSerArg 343
|||||:|||||:|||||: |||: |||
975 CCACTTAACATATGAGACTCCAGATGTTTAACTTCCGTGCAGACAAA 1024
|||||:|||||:|||||: |||: |||
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
|||||:|||||:|||||: |||: |||
1025 GGTGTGTATGACGCTGCATCTCATGTTCGCAAGTGTCTTGGCCCT 1074
|||||:|||||:|||||: |||: |||
360 LeuSerValAlaAlaAspAlaLeuSerAsnPheTyTyProProVal 376
|||||:|||||:|||||: |||: |||
1075 CTTCTGATGATGCTGCAAAATGCACTCAAAACATATATACCAACAGT 1124
|||||:|||||:|||||: |||: |||
376 lGlyAlaValThrIleSerTyTyProGluGluAlaIleArgSpGluArg 393
|||||:|||||:|||||: |||: |||
```

```
1125 TGCACGATCTATCTCTGTAACCGGAAAGCAATCCGAAGAATGTT 1174
393 euvalaspelgyluleuylsglypneqlygluleuhsiproahgthgln 409
1175 TGATAGATGGGAACTAAGAGGTTTGGCAATGTCACACGACGACAA 1224
410 gllvalgluThrleuaglyThrleTySerSerSerleupheProanar 426
1225 GGAATGAAACATTAGGACTATATTACAGCTCCCTCATCTTTCCAAATCG 1274
426 galaprolgslgylarvalleuLeuLeuasnTyrlleglylvalalysa 443
1275 CGCACCGCGCGGAGAAATTTGCTGTGAATACATTGGCGGCTACAA 1324
443 snpTrogliileuSerlyThrleuSerGlnleuValgluValasp 459
1325 ACACCGGAATTCGTCCAGCTCAAGGTGAGTGTAGGAGAGGTTGAC 1374
460 ArgaspleuarqlysmetleuileysProlysaliaglnasproleuva 476
1375 AGAGATTGGAGAAATGCTAATTATTAAGCCTTAATTCGACCCGACTTAA 1424
476 lvalglvalarvalTrrProglinalaleProglInpheleuValglu 493
1425 AATGAGATGAGGTATGGGCTCAAGCATTCCTCAGTTCTTACTGTGTC 1474
493 lsleuasprThleuSerThrAlalysalaAlaMetasnAspAsnglyeu 509
1475 ACTTGATATCTCTGACACGCGCTAATCATCTCAACGCTTCGCGGTAC 1524
510 gluglyleupheleuGlYlYAsnTyValSerGlyValalaleuGlYar 526
1525 GAGGGCTATTTTGGTGCAATTACGCTCGCTGTAGCCTTAGGCGG 1574
526 gCysValgluGlYalatyrgluValAlaSerGluValThrglypheuS 543
1575 GTGTGTAAGAGCGCATATGAACCGCATGAGGTCAACAACCTCATGT 1624
543 eArqTyrAlaTyrlYs 548
1625 CACGCTACGCTTACAG 1641

seq_name: /cgn2_6/prodata/1/lna/68_COMB.seq:US-09-071-296-1

seq documentation block:
; Sequence 1, Application US/09071296
; Patent No. 6177245
; GENERAL INFORMATION:
; APPLICANT: Ward, Eric R
; APPLICANT: Voltrath, Sandra
; TITLE OF INVENTION: Manipulation of Protoporphyrinogen
; TITLE OF INVENTION: Oxidase Enzyme Activity in Eukaryotic Organisms
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,296
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/261,198
; FILING DATE: 16-JUN-94
; ATTORNEY/AGENT INFORMATION:
```

```
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1748/CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1719 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 31..1644
OTHER INFORMATION:
OTHER INFORMATION: sequence from pMDC-2"
US-09-071-296-1
```

```
alignment_scores:
Quality: 2077.00 Length: 539
Ratio: 4.230 Gaps: 4
Percent Similarity: 91.095 Percent Identity: 75.139
```

```
alignment_block:
US-09-508-418-2 x US-09-071-296-1 ..
```

```
Align seg 1/1 to: US-09-071-296-1 from: 1 to: 1719
```

```
19 SerSerProleuAlapheleuAsnArgThrSerPheIleProPheSerse 35
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
28 TCATGAGATTAATCTCTCTCCGCGAGACTCAATCCCTCTTCGCCGTC 77
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
35 rIleSerlySArgAsnSerValAsnCysAsnglyTrp.....ArgT 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
78 GTTTGGAAGCCCAAT...CTCCGATTAATGTTTATTAAGCCTTATAGAC 124
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
49 hArgCysSerValAlaLysAspTyrThrValProSerSerAlaValasp 65
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 TCCGTTGTCAGTGGCGGTGACCAACCGCTGCGATCTTCAAAATCGAA 174
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 gLYglYProAlaAlaGluLeu.....AspCysValIleValglYAlaGl 80
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
175 GCGGAGAGGAGCACACACATCAGCAGGANTGTGTATGTGTGCGCGGAGS 224
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 yIleSerGlyLeuCysIleAlaGlnValMetSerAlaAsnTyf..... 94
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
225 TATTAGGCTCTTGGCATCGCTCAGCGGCTTCTACTAAGCACTCGATG 274
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
95 ....ProAsnleuMetValThrGluAlaArgAspArgAlaGlyIlyasn 109
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
275 CTGCTCCGAATTAATGTGTACCGAGGCTAAGATGTGTGTGGAGGCAAC 324
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
110 IleThrValAlaGluArgAspGlyTyrLeuTrrPgluGluGlyProAsnse 126
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
325 ATTATCATCTGCTGAAGAGATGTTTCTCTGGAAGAAGGTCCTCAATAG 374
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 rPheGlnProSerAspProMetleuThrMetAlaValaspCysGlyleu 143
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
375 TTTTCAACCGCTGATCTATGCTCACTAAGGCTAGAGTAGAGGTTTGA 424
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
143 ysAspAspleuValleuGlYAspProAsnAlaProArgPheValleuTrrp 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
425 AGGATGATTTGGTGGAGATCTACTGCGCAAGGTTTGTGTGTGG 474
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
160 LysGlyLysLeuArgProValProSerLysLeuThrAspleuProhepb 176
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
475 AATGGAAATTTGAGCGCGTTCCATCGAAGCTTAACGACTTACGTTCTT 524
```

176 eAspleuMetSerIleProGlyLysLeuArgAlaGlyPheGlyProIleG 193
|||||.....
525 TGATTTGATGAGTATGTTGGGAGATTAGAGCTGTTGGTTCACCTTG 574
193 LLeuArgProSerProGlyHisGluGluSerValGluGlnPheVal 209
|||||.....
575 GCATTGACCGCTCACCTCCAGGTGTGTAAGATTTGTGGAGAGATTGTA 624
210 ArgArgAsnLeuGlyGlyGluValPheGluArgLeuIleGluProPheCy 226
|||||.....
625 CGGGGTAACTCGGTGATGAGGTTTGGAGCGCTGATTGACCGTTTG 674
226 sSerGlyValTyrValGlyAspProSerLysLeuSerMetLysAlaAlaP 243
|||||.....
675 TTCAGGTTTATGCTGTATCCTTCAAAAGTGAAGATGAAAGCACCGT 724
243 heGlyLysValTyrPheLysLeuGluGlnThrGlySerIleIleGly 259
|||||.....
725 TTGGGAAGGTTTGAAACTAGACAAATGTTGGAACATATATAGTGT 774
260 ThrPheLysAlaAlaIleLysGluArgSerSerThrProLysAlaProArg 276
|||||.....
775 ACTTTTAAAGCAATTCAAGAGAGGAAAAAGCTCCCAAGCAGAACGAGA 824
276 ProArgLeuProLysGlyGlnThrValGlySerPheArgLysG 293
|||||.....
825 CCGCGCGCTGCAAAACCAACGAGGCAAAAGTTGTTCTTTCAGGAAG 874
293 LLeuArgMetLeuProAspAlaIleSerAlaArgLeuGlySerLysLeu 309
|||||.....
875 GACTTCGAATGTTGCCAGAGCAATATCTGCAGATAGTAGCAAAAGT 924
310 LysLeuSerTyrPheLysLeuSerIleThrLysSerGluLysGlyTyr 326
|||||.....
925 AAGTTGCTTGGAGCTCTCAGGATCATCAAGCTGAGAGCGAGAGATA 974
326 rHisLeuThrTyrGluThrProGluGlyValValSerLeuGlnSerArg 343
|||||.....
975 CAACTTAAATATGAGAGCTCAAGTGTGTTAGTTTCCTGACAGCAAAA 1024
343 erIleValMetThrValProSerTyrValAlaSerAsnIleLeuArgPro 359
|||||.....
1025 GTGTGTAATGACGGTGCATCTCATCTGCAAGTGTCTTGGCGCCT 1074
360 LeuSerValAlaAlaAlaAspAlaLeuSerAspPheTyrTyrProProVa 376
|||||.....
1075 CTTTGTGAATCTCTGCATAATGCACTCTCAAAACTATATATACCCACAGT 1124
376 lGlyAlaValThrIleSerTyrProGlnGluAlaIleArgAspGluArgL 393
|||||.....
1125 TGCAGCAGTATCTATCTCGTACCCGAAAGCAATCCGAAACGAATGTT 1174
393 euValAspGlyGluLeuLysGlyPheGlyGlnLeuHisProArgThrGln 409
|||||.....
1175 TGATAGATGTTGAACATAAGGTTTGGGCAATTCATCCAGCAGCA 1224
410 GlyValGluThrLeuGlyThrIleTyrSerSerSerLeuPheProAsnAr 426
|||||.....
1225 GGAGTTGAAACATTAGCAATATCTACAGTCTCTCACTTTCCAAATCG 1274
426 gAlaProLysGlyArgValLeuLeuLeuAsnTyrIleGlyValAlaLysA 443
|||||.....
1275 CGGACCGCCCGGAAAGATTTTGTGTGTAACATACATTGGGGGTTCACA 1324
443 snProGluIleLeuSerLysThrGluSerGlnLeuValGluValAsp 459
|||||.....
1325 ACACCGGAATTTCTGCAAGTCTGAAGGTGAGTTACTGGAAGAGTTGAC 1374
460 ArgAspLeuArgLysMetLeuIleLysProLysAlaGlnAspProLeuVa 476
|||||.....
1375 AGAGATTGAGGAAATGCTAATTAAGCCTAATTCACACGATCCACTTAA 1424
476 lValGlyValArgValTyrProGlnAlaIleProGlnPheLeuValGlyH 493

seq_name: /cgn2_6/ptodata/1/ina/5B_COMB.seq:US-08-808-931-11
seq_documentation_block:
; Sequence 11 Application US/08808931
; Patent No. 5939602
; GENERAL INFORMATION:
; APPLICANT: Volrath, Sandra
; APPLICANT: Johnson, Marie
; APPLICANT: Potter, Sharon
; APPLICANT: Ward, Eric
; APPLICANT: Helfetz, Peter
; TITLE OF INVENTION: DNA Molecules Encoding Plant
; TITLE OF INVENTION: Protochlorophyllinogen oxidase and inhibitor-Resistant Mutants
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5939602artis Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,931
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,705
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/013,612
; FILING DATE: 28-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/020,003
; FILING DATE: 21-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: CGC 1847
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

```

?
? HYPOTHETICAL: NO
? ORIGINAL SOURCE:
? ORGANISM: soybean
? IMMEDIATE SOURCE:
? CLONE: PMDC-12 (NRRL B-21516)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 55..1683
? OTHER INFORMATION: /product="soybean protox-1"
US-08-808-931-11

alignment_scores:
  Quality: 2062.50      Length: 550
  Ratio: 4.218          Gaps: 7
  Percent Similarity: 88.909   Percent Identity: 73.273

alignment_block:
  US-09-508-418-2 x US-08-808-931-11  ..

Align seg 1/1  to: US-08-808-931-11  from: 1  to: 1847

   3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSer 19
   ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
102 TCTTCGCCCTCCCTCCATTCGCCACCTCTTCTTCCACTCTCCACTC 151
   ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
   19 rSerProLeuAlaPheLeuAsnArgThrSerPheIleProPheSer 36
   ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
152 GAAA.....ATTCCC.....T 162

   36 LeSerIlyAsnSerValAsnCysAsnGlyTrpArgThrArgCysSe 52
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
163 CGCTCTGCCCTAACCTATT.....CTACCTCTCTC 194

   52 rValAlaIlyAspTyrThrVal...ProSerSerAlaValAspGly 68
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
195 CATTCGGGAGAAATCCACCGCGCTCCGCCCAAAACAGACATCCGCC 244

   68 rValAlaIleGluLeuAspCysValIleValGlyAlaGlyIleSer 84
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 CC.....GTGGACTGCTGCTGCTGGGAGGCGTACGGGCTC 285

   85 CysIleIleGlnValMetSerAlaAsnTyr.....ProAsnLeuMet 99
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
286 TGCATCGCCCAAGCCCTCGCCACCAACACGCCCATGCAACGTCGT 335

   99 lThrGluAlaArgAspArgAlaGlyGlyAsnIleThrThrValGlu 116
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 CACGGAGGCCCGAGACCGCTCGCGCGCACATCACACGATGAGAGG 385

116 spGlyTyrLeuTrpGluGluGluProAsnSerPheGlnProSerAs 132
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
386 ACGGATACCTCTGGGAGAGAGGCCACACACCTTCCACCTTGTGATCA 435

133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuVal 149
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 ATGCTCACCATGCTGTGTGACAGAGGTTAAAGATGAGCTGTTGGG 485

149 yAspProAsnAlaProArgPheValLeuTrpLysGlyLysLeuArg 166
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
486 GGATCTCATGACCTCGCTTGTGTGTGTGAGACAGAAAGTTGAGCGG 535

166 aProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIle 182
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 TGCCGGGAGAGCTGACGTGATTTGCTTTCTTTCATCTGATGACGAT 585

183 GlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProSer 199
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
586 GGCAGAAATCAGAGCGTGTGTGTGTGTGTGAAATTCGCTCCCTCC 635

199 oGlyHisGluSerValGluGlnPheValArgArgAsnLeuGlyGly 216
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
636 AGTCATAGAGAAATCGTGTGAAGTTGTTCGTGCAACCTTGATG 685
```

```

216 lValAlaPheGluArgLeuIleGluProPheCysSerGlyValTyrVal 232
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
686 AGGTTTTCAGCGTTGATGATGACCTTTTGTTCAGGGCTGTATGCA 735

233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTyrP 249
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
736 GATCTCTTCAAAATTAAGTATGAAACACACATTCGGGAAAGTTGA 785

249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIle 266
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
786 GAAAAAATGCTGATGATATGTTGAGCACTTCAAGCAATATCAAG 835

266 lArgSerSerThrProLysAlaProArgAspProArgLeuProLys 282
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
836 AGAGAAATGAGCTTCAAAACCACTCGAGATCCGCTCTGCCAAAACA 885

283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeu 299
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
886 AAGGTCACACTGTGGATCTTCCGGAAGGACCTTACCATGTGCTGCA 935

299 pAlaIleSerAlaArgLeuGlySerLysLeuLysLeuSerTrpLys 316
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
936 TCGAATTTCTGCCAGACTGAGCAACAAGTAAATATCTTGAAAGCT 985

316 eSerIleThrLysSerGluLysGlyLysGlyTyrPheIleThrTyr 332
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
986 CAGATATGTGTAACCTGATGATGAGAGTACAGTTTGACATATGAACA 1035

333 ProGluGlyValValSerLeuGlnSerArgSerIleValMetThr 349
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1036 CCAGAGAGAGTGGTTCCTTTCGACGTCAAAACCTGTCCTGACCAT 1085

349 oSerTyrValAlaIleAsnIleLeuArgProLeuSerValAlaAla 366
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1086 TTCCTATGTGTGATGACATTCCTGCTGCTGCTGCTGCTGCTGCA 1135

366 spAlaLeuSerAsnPheTyrTyrProProValGlyAlaValThr 382
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1136 ATGCATTTTCAAAAGTTTATTTACCTCCAGTTGCTGACATTTCCAT 1185

383 TyrProGluGlnAlaIleArgAspGluArgLeuValAspGlyGlu 399
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1186 TATCCAAAAGAGCTATGTAGATCAGAAATCTGATGAGTGAAGTGA 1235

399 sGlyPheGlyGlnLeuHisProArgThrGlnGlyValGluThrLeu 416
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1236 GGGGTTTGTCAATTCATTCACCTGACGTAGCCAGAGAGTGAACAT 1285

416 hTrpIleYSerSerSerLeuPheProAsnArgAlaProLysGly 432
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1286 CTATATACAGCTCATCATTTCCCAACCGACACACCTGGAAGGCT 1335

433 LeuLeuLeuAsnTyrIleGlyAlaLysAsnProGluIleLeuSer 449
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1336 CTACTCTTGATTAATCATTTGAGAGGCACTAATCTGGAATTTTAT 1385

449 sThrGluSerGlnLeuValGlyValAlaAspArgAspLeuArgLys 466
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1386 GACGGACAGTAACTGTGAAACACTTGATGCAGATTTGAGGAAATCC 1435

466 euIleLysProLysAlaGlnAspProLeuValAlaGlyValArg 482
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1436 TTATTAACCAAAATGCCAGAGATCCATTTGTAGTGGGCTGAGACT 1485

483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeu 499
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1486 CCTCAAGCTATTCACAGTTTATGTTGCCATTTGATCTTCTTATAG 1535

499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGluGlyLeuPheLeu 516
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1536 TGCTAAAGCTTCTATCAGAAATACCTGGTTTGAAGGCTTCTTCC 1585
```

```

516  lysnfyrrvalserglvalalaleucliyarqcyvalglunglvalaary 532
      |||||
1586  gtaattatgctgcgtggcttgccttggagacatcgccttgagaggacctat 1633
      |||||
533  gluvallasericvalthrnglyphelueserarglyrralalyrrlys 548
      |||||
1636  gaggtgacgacctaaagtaaacgattttctctcacaaatagaggttcaaaa 1683
      |||||
seq_name: /cgn2_6/ptdata/1/lna/6A.COMB.seq:US-08-808-323-11

```

```

seq documentation block:
: Sequence 11, Application US/08808323
: Patent No. 6018105
:
: GENERAL INFORMATION:
: APPLICANT: Johnson, Marie
: APPLICANT: Volrath, Sandra
: APPLICANT: Ward, Eric
: TITLE OF INVENTION: Promoters from Plant
: TITLE OF INVENTION: Protoporphylinogen Oxidase Genes
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: No. 6018105artis Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591-9005
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/808,323
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/012,705
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/013,612
: FILING DATE: 28-FEB-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/020,003
: FILING DATE: 21-JUN-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Weigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: CGC 1846
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-6689
:
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1847 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHEICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: soybean
: IMMEDIATE SOURCE:
: CLONE: pMDC-12 (NRRL B-21516)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 55..1683
: OTHER INFORMATION: /product= "soybean protox-1"
US-08-808-323-11

```

```
alignment_scores:
    Quality: 2062.50
    Length: 550
```

| | | | |
|---------------------|----------------------|-------------------|----------|
| Ratio: | 4.218 | Gaps: | 7 |
| Percent Similarity: | 88.909 | Percent Identity: | 73.273 |
| alignment block: | | | |
| US-09-508-418-2 | x | US-08-808-323-11 | .. |
| Align seg 1/1 | to: US-08-808-323-11 | from: 1 | to: 1847 |

3 ThrTrpProIleAlaAsnHisProAsnIlePheThrHisGlnSerSerSe 19
 102 TCTTGGCCCTCCCTCCATTCGCCCAACCTCTTCTTCACTCCCTCCACTC 151
 19 rSerProLeuAlaPheLeuAsnArGthrSerPheIleProPheSerSerI 36
 152 GAAA.....ATTCCTC.....T 162
 163 GCCTCTGCCCTTAACCTTAT.....CTACCTCTCTC 194
 52 rValAlaLysAspThrVal...ProSerSerAlaValAspGlyGly 68
 195 CATTCGCGAGAAATTCACACCGCGCTCTCCGCCAAACAGACATCTCCGCC 244
 68 rValAlaIleGluLeuAspCysValIleValGlyAlaGlyIleSerGlyLeu 84
 245 CC.....GTGGACTGGCTGCTGCTGGCGGAGGGGCTCAGCGGCCTC 285
 85 CysIleLeuGlnValMetSerAlaAsnTrp.....ProAsnLeuMetVal 99
 286 TGCATTCGCCACGAGCCCTCCGACACCAACCGCCAAATGCCAACGTCGTCGT 335
 99 LThrGluAlaArgAspArgAlaGlyGlyAsnIleThrValIleGluArg 116
 336 CACGAGAGCCCGAGACCGCGCTGGCGCGCAACATCAACACGATGGAGAGGG 385
 116 spGlyTrpLeuTrpGluGluGlyProAsnSerPheGlnProSerAspPro 132
 386 ACGGATTCCTCTGTGGAAAGAGCCCAACACAGCTTCACCTTCGTGATCCA 435
 133 MetLeuThrMetAlaValAspCysGlyLeuLysAspAspLeuValLeuGln 149
 436 ATGCTCAACATGGTGGTGGACAGTGGTTAAAGATGAGCTGTGTTTGGG 485
 149 yAspProAsnAlaProArgPheValLeuTrpLysLeuArgProVal 166
 486 GGATCCTCGATGCACCTCGGTTGTGTGTGGACAAGAAATTGAGGCCGG 535
 166 alProSerLysLeuThrAspLeuProPhePheAspLeuMetSerIlePro 182
 536 TGCCCGGAGAGCTGACTGATTTGGCTTCTTGTGACTGATGAGCATTTGGT 585
 183 GlyLysLeuArgLalaIlyPheGlyProIleGlyLeuArgProSerProPr 199
 586 GGCAGAAATCGAGCGCTGGCTTGTGGCTGGAAATGGGCTCTCTCTCC 635
 199 OGlyHisGluLysLeuValGlnGlnPheValAlaArgAsnLeuGlyGly 216
 636 AGGCGATGAGGAATCGGTTGGAAGACTTGTCTTCGCGAGACCTTGGTGATG 685
 216 LuValPheGluArgLeuIleGlnProPheCysSerGlyValIlyValGly 232
 686 AGGTTTTTGAACGGTTGATAGAGCTTTTGTTCAGGGGCTATGACAGGC 735
 233 AspProSerLysLeuSerMetLysAlaAlaPheGlyLysValTrpLysLe 249
 736 GATTCCTTCAAAATTAAGTATGAACACACATTCGGGAAATTTGGAGAGCT 785
 249 uGluGluThrGlyGlySerIleIleGlyGlyThrPheLysAlaIleLysG 266
 786 GGAAGAAATATGGTGGTACATTATTGTGGAGAACCTTTAAAGCAATATCAAG 835
 266 LuArgSerSerThrProLysAlaProArgAspProArgLeuProLysPro 282

```

|||||:..... ||| |||||||||||||||||||
836 AGAAGATGAGCTCCAAACCACTCGAGATCCGGCTCGCAAAACCA 885
283 LysGlyGlnThrValGlySerPheArgLysGlyLeuArgMetLeuProAs 299
|||||:..... ||| |||||||||||||||||||
886 AAGGCTAGACGTGTGGACTTTCCGGAAGGCACTTACCATGTTGCTTA 935
299 palalieserialaargleuglyserlysleuylsleusertrpulsleus 316
|||||:..... ||| |||||||||||||||||||
936 TCGAATTTCCGACAGCTAGGCAACAAAGTAAGTATCTTGGAACCTTT 985
316 eiserlietherlysleuglyglylytyrhisleuhyrtyrlyglnthr 332
|||||:..... ||| |||||||||||||||||||
986 CAAGTATTAGTAACTGAGTAGTGAGAGACGTTTGACATATGTAAACA 1035
333 ProGluGlyValValSerLeuGlnSerArgSerLeuValMetThrValPr 349
|||||:..... ||| |||||||||||||||||||
1036 CCGAAGAGAGTGTCTTCTTGCGAGTCGCAAAAGTGTCTCTGACCATTC 1085
349 oSerTyrrValAlaSerAsnIleLeuArgProLeuSerValAlaAlaAla 366
|||||:..... ||| |||||||||||||||||||
1086 TTCCTATGTGTAGTACATGCTCGCTCCTCTGCTGCTGCTGCTGCA 1135
366 spalaleuserasphenetrrtyrproprovalalvalalthrileser 382
|||||:..... ||| |||||||||||||||||||
1136 ATGCACTTTAAAGTTTATTATACCTCCACTGCTGCTGCACTTCCATATCC 1185
383 TyrProGlnLualalalaraspGluArgLeuValaspGlyLuleuLy 399
|||||:..... ||| |||||||||||||||||||
1186 TATCCAAAGACGCTATTAGATCAGAACTGCTATAGATGTGAGTTGAA 1235
399 sglyPheGlyGlnLeuHisProArgThrGlnGlyValGlnThrLeuGlyT 416
|||||:..... ||| |||||||||||||||||||
1236 GGGGTTTGCTCAATTCATCCAGTACCCAGAGAGTGAACATTTAGGAA 1285
416 hriletyrSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
|||||:..... ||| |||||||||||||||||||
1286 CTATATACAGCTCATCTACTATCCCAACGAGACACCACTGGAAGGTT 1335
433 LeuLeuLeuAsnTyrrileGlyGlyAlaLysAsnProGluIleLeuSerLy 449
|||||:..... ||| |||||||||||||||||||
1336 CTACTCTTGAATTCATTGAGAGGCACTAAATACCTGGAATTTTATCGAA 1385
449 sThrGluSerGlnLeuValGluValAlaAspArgAspLeuArgLysMetL 466
|||||:..... ||| |||||||||||||||||||
1386 GACGGACAGTGAACCTGTGGAAACAGTTGATCCGAGATTGAGGAAATCC 1435
466 euileLysProLysAlaGlnAspProLeuValValGlyValArgValTyrP 482
|||||:..... ||| |||||||||||||||||||
1436 TTATAAACCCAAATGCCAGAGATCCATTGTAGTGGGGTGAGACTGTGG 1485
483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSerTh 499
|||||:..... ||| |||||||||||||||||||
1486 CCTCAGCTATTCACAGTTCTTGTAGTGGCCATCTTGATCTTCAAGATGT 1535
499 rAlaLysAlaAlaMetAsnAspAsnGlyLeuGlnGlyLeuPheLeuGlyG 516
|||||:..... ||| |||||||||||||||||||
1536 TGCCTAAAGCTCTATACAAATACTGGGTTGAAAGGCTCTCTCTTGGGG 1585
516 LysAsnTyrrValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTyr 532
|||||:..... ||| |||||||||||||||||||
1586 GTAATTATAGTCTGTGGTGTCTCTTGGGAGAGATGCGTTGAGGAGCCTAT 1635
533 GluValAlaAspGluValThrGlyPheLeuSerArgTyrrAlaTyrrLys 548
|||||:..... ||| |||||||||||||||||||
1636 GAGGTAGACGCTGAAGTAACGATTCTTCACAAATAGAGTGTCAAA 1683

```

seq_name: /cgn2_6/prodata/1/ina/6A_COMB.seq:US-09-050-603A-11

```

seq_documentation_block:
; Sequence 11, Application US/09050603A
; Patent No. 6023012
; GENERAL INFORMATION:

```

```

? APPLICANT: Volrath, Sandra
? APPLICANT: Johnson, Marie
? APPLICANT: Potter, Sharon
? APPLICANT: Ward, Eric
? APPLICANT: Helfetz, Peter
? TITLE OF INVENTION: DNA Molecules Encoding Plant
? TITLE OF INVENTION: Protoporphyrinogen Oxidase
? NUMBER OF SEQUENCES: 37
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: No. 6023012artis Corporation
? STREET: 3054 Cornwallis Road
? CITY: Research Triangle Park
? STATE: NC
? COUNTRY: USA
? ZIP: 27709
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/050,603A
? FILING DATE: 30-MAR-1998
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/808,931
? FILING DATE: 28-FEB-1997
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/012,705
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/013,612
? FILING DATE: 28-FEB-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/020,003
? FILING DATE: 21-JUN-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Melgs, J. Timothy
? REGISTRATION NUMBER: 38,241
? REFERENCE/DOCKET NUMBER: CGC 1847
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (919) 541-8587
? TELEFAX: (919) 541-8689
? INFORMATION FOR SEQ ID NO: 11:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1847 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? HYPOTHETICAL: NO
? ORIGINAL SOURCE:
? ORGANISM: soybean
? IMMEDIATE SOURCE:
? CLONE: pMDC-12 (NRRL B-21516)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 55..1683
? OTHER INFORMATION: /product= "soybean protox-1"
US-09-050-603A-11

```

```

alignment_scores:
    Quality: 2062.50      Length: 550
    Ratio: 4.218          Gaps: 7
    Percent Similarity: 88.909    Percent Identity: 73.273

```

alignment_block: US-09-508-418-2 x US-09-050-603A-11 ..

Align seg 1/1 to: US-09-050-603A-11 from: 1 to: 1847

3 ThrThrProIleAlaAsnHisProAsnIlePheThrHisGlnSerSer 16


```
102 TTTGGCCCCCTCCCAATTCCTCCCAACCTTTCTTCCACTCTCCACTC 151
103 :
104 :
105 :
106 :
107 :
108 :
109 :
110 :
111 :
112 :
113 :
114 :
115 :
116 :
117 :
118 :
119 :
120 :
121 :
122 :
123 :
124 :
125 :
126 :
127 :
128 :
129 :
130 :
131 :
132 :
133 :
134 :
135 :
136 :
137 :
138 :
139 :
140 :
141 :
142 :
143 :
144 :
145 :
146 :
147 :
148 :
149 :
150 :
151 :
152 :
153 :
154 :
155 :
156 :
157 :
158 :
159 :
160 :
161 :
162 :
163 :
164 :
165 :
166 :
167 :
168 :
169 :
170 :
171 :
172 :
173 :
174 :
175 :
176 :
177 :
178 :
179 :
180 :
181 :
182 :
183 :
184 :
185 :
186 :
187 :
188 :
189 :
190 :
191 :
192 :
193 :
194 :
195 :
196 :
197 :
198 :
199 :
200 :
201 :
202 :
203 :
204 :
205 :
206 :
207 :
208 :
209 :
210 :
211 :
212 :
213 :
214 :
215 :
216 :
217 :
218 :
219 :
220 :
221 :
222 :
223 :
224 :
225 :
226 :
227 :
228 :
229 :
230 :
231 :
232 :
233 :
234 :
235 :
236 :
237 :
238 :
239 :
240 :
241 :
242 :
243 :
244 :
245 :
246 :
247 :
248 :
249 :
250 :
251 :
252 :
253 :
254 :
255 :
256 :
257 :
258 :
259 :
260 :
261 :
262 :
263 :
264 :
265 :
266 :
267 :
268 :
269 :
270 :
271 :
272 :
273 :
274 :
275 :
276 :
277 :
278 :
279 :
280 :
281 :
282 :
283 :
284 :
285 :
286 :
287 :
288 :
289 :
290 :
291 :
292 :
293 :
294 :
295 :
296 :
297 :
298 :
299 :
300 :
301 :
302 :
303 :
304 :
305 :
306 :
307 :
308 :
309 :
310 :
311 :
312 :
313 :
314 :
315 :
316 :
317 :
318 :
319 :
320 :
321 :
322 :
323 :
324 :
325 :
326 :
327 :
328 :
329 :
330 :
331 :
332 :
333 :
334 :
335 :
336 :
337 :
338 :
339 :
340 :
341 :
342 :
343 :
344 :
345 :
346 :
347 :
348 :
349 :
350 :
351 :
352 :
353 :
354 :
355 :
356 :
357 :
358 :
359 :
360 :
361 :
362 :
363 :
364 :
365 :
366 :
367 :
368 :
369 :
370 :
371 :
372 :
373 :
374 :
375 :
376 :
377 :
378 :
379 :
380 :
381 :
382 :
383 :
384 :
385 :
386 :
387 :
388 :
389 :
390 :
391 :
392 :
393 :
394 :
395 :
396 :
397 :
398 :
399 :
400 :
401 :
402 :
403 :
404 :
405 :
406 :
407 :
408 :
409 :
410 :
411 :
412 :
413 :
414 :
415 :
416 :
417 :
418 :
419 :
420 :
421 :
422 :
423 :
424 :
425 :
426 :
427 :
428 :
429 :
430 :
431 :
432 :
433 :
434 :
435 :
436 :
437 :
438 :
439 :
440 :
441 :
442 :
443 :
444 :
445 :
446 :
447 :
448 :
449 :
450 :
451 :
452 :
453 :
454 :
455 :
456 :
457 :
458 :
459 :
460 :
461 :
462 :
463 :
464 :
465 :
466 :
467 :
468 :
469 :
470 :
471 :
472 :
473 :
474 :
475 :
476 :
477 :
478 :
479 :
480 :
481 :
482 :
483 :
484 :
485 :
486 :
487 :
488 :
489 :
490 :
491 :
492 :
493 :
494 :
495 :
496 :
497 :
498 :
499 :
500 :
501 :
502 :
503 :
504 :
505 :
506 :
507 :
508 :
509 :
510 :
511 :
512 :
513 :
514 :
515 :
516 :
517 :
518 :
519 :
520 :
521 :
522 :
523 :
524 :
525 :
526 :
527 :
528 :
529 :
530 :
531 :
532 :
533 :
534 :
535 :
536 :
537 :
538 :
539 :
540 :
541 :
542 :
543 :
544 :
545 :
546 :
547 :
548 :
549 :
550 :
551 :
552 :
553 :
554 :
555 :
556 :
557 :
558 :
559 :
560 :
561 :
562 :
563 :
564 :
565 :
566 :
567 :
568 :
569 :
570 :
571 :
572 :
573 :
574 :
575 :
576 :
577 :
578 :
579 :
580 :
581 :
582 :
583 :
584 :
585 :
586 :
587 :
588 :
589 :
590 :
591 :
592 :
593 :
594 :
595 :
596 :
597 :
598 :
599 :
600 :
601 :
602 :
603 :
604 :
605 :
606 :
607 :
608 :
609 :
610 :
611 :
612 :
613 :
614 :
615 :
616 :
617 :
618 :
619 :
620 :
621 :
622 :
623 :
624 :
625 :
626 :
627 :
628 :
629 :
630 :
631 :
632 :
633 :
634 :
635 :
636 :
637 :
638 :
639 :
640 :
641 :
642 :
643 :
644 :
645 :
646 :
647 :
648 :
649 :
650 :
651 :
652 :
653 :
654 :
655 :
656 :
657 :
658 :
659 :
660 :
661 :
662 :
663 :
664 :
665 :
666 :
667 :
668 :
669 :
670 :
671 :
672 :
673 :
674 :
675 :
676 :
677 :
678 :
679 :
680 :
681 :
682 :
683 :
684 :
685 :
686 :
687 :
688 :
689 :
690 :
691 :
692 :
693 :
694 :
695 :
696 :
697 :
698 :
699 :
700 :
701 :
702 :
703 :
704 :
705 :
706 :
707 :
708 :
709 :
710 :
711 :
712 :
713 :
714 :
715 :
716 :
717 :
718 :
719 :
720 :
721 :
722 :
723 :
724 :
725 :
726 :
727 :
728 :
729 :
730 :
731 :
732 :
733 :
734 :
735 :
736 :
737 :
738 :
739 :
740 :
741 :
742 :
743 :
744 :
745 :
746 :
747 :
748 :
749 :
750 :
751 :
752 :
753 :
754 :
755 :
756 :
757 :
758 :
759 :
760 :
761 :
762 :
763 :
764 :
765 :
766 :
767 :
768 :
769 :
770 :
771 :
772 :
773 :
774 :
775 :
776 :
777 :
778 :
779 :
780 :
781 :
782 :
783 :
784 :
785 :
786 :
787 :
788 :
789 :
790 :
791 :
792 :
793 :
794 :
795 :
796 :
797 :
798 :
799 :
800 :
801 :
802 :
803 :
804 :
805 :
806 :
807 :
808 :
809 :
810 :
811 :
812 :
813 :
814 :
815 :
816 :
817 :
818 :
819 :
820 :
821 :
822 :
823 :
824 :
825 :
826 :
827 :
828 :
829 :
830 :
831 :
832 :
833 :
834 :
835 :
836 :
837 :
838 :
839 :
840 :
841 :
842 :
843 :
844 :
845 :
846 :
847 :
848 :
849 :
850 :
851 :
852 :
853 :
854 :
855 :
856 :
857 :
858 :
859 :
860 :
861 :
862 :
863 :
864 :
865 :
866 :
867 :
868 :
869 :
870 :
871 :
872 :
873 :
874 :
875 :
876 :
877 :
878 :
879 :
880 :
881 :
882 :
883 :
884 :
885 :
886 :
887 :
888 :
889 :
890 :
891 :
892 :
893 :
894 :
895 :
896 :
897 :
898 :
899 :
900 :
901 :
902 :
903 :
904 :
905 :
906 :
907 :
908 :
909 :
910 :
911 :
912 :
913 :
914 :
915 :
916 :
917 :
918 :
919 :
920 :
921 :
922 :
923 :
924 :
925 :
926 :
927 :
928 :
929 :
930 :
931 :
932 :
933 :
934 :
935 :
936 :
937 :
938 :
939 :
940 :
941 :
942 :
943 :
944 :
945 :
946 :
947 :
948 :
949 :
950 :
951 :
952 :
953 :
954 :
955 :
956 :
957 :
958 :
959 :
960 :
961 :
962 :
963 :
964 :
965 :
966 :
967 :
968 :
969 :
970 :
971 :
972 :
973 :
974 :
975 :
976 :
977 :
978 :
979 :
980 :
981 :
982 :
983 :
984 :
985 :
986 :
987 :
988 :
989 :
990 :
991 :
992 :
993 :
994 :
995 :
996 :
997 :
998 :
999 :
1000 :
```

```
seq_name: /cgn2_6/pdata/./ina/6a_COMB.seq.us-09-102-420B-11
seq_documentation_block:
: Sequence 11, Application US/09102420B
: Patent No. 6084155
: GENERAL INFORMATION:
: APPLICANT: Voltrath, Sandra
: APPLICANT: Johnson, Marie
: APPLICANT: Ward, Eric
: APPLICANT: Helfetz, Peter
: TITLE OF INVENTION: HERBICIDE-TOLERANT PROTOPORPHYRINGEN
: TITLE OF INVENTION: OXIDASE ("PROTOX")
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NO. 6084155artis Corporation
```



```

299 pAlaIleSerAlaArgLeuGlySerIysLeuLysSerTrpIysLeuS 316
|||||
936 TCCAATTTCGCACACTAGCAGCAACAAAGTAAAGTTATCTTGAGAGCTT 985
316 eSerIleThrIysSerGluIysGlyGlyTrpHisLeuThrTrpGluThr 332
|||||
966 CAGATATTAGTAAGTACTGATAGTGTGAGAGTTCACATTGTGACATATGAACA 1035
333 ProGluGlyAlaValSerLeuGlnSerArgSerIleValMetThrValPr 349
|||||
1036 CCAGAGAGAGTGGTTCTTCTGACAGCAAAAGCTGTCTCCGACCATTC 1085
349 oSerTrpAlaIleSerAsnIleLeuArgProLeuSerValAlaIleAla 366
|||||
1086 TTCCTAGTGTCTAGTACATTGCTGCTCTCTCTGCTGCTGCTCAG 1135
366 sPAlaIleSerAsnPhetTrpProProValGlyAlaValThrIleSer 382
|||||
1136 ATGCACTTTCAAAGTTTATTACCTCCAGTTCGTCGATTCATATCC 1185
383 TrpProGlnGluAlaIleArgAspGluArgLeuValAspGlyGlyLeu 399
|||||
1186 TATCCAAAGAGAGCTATTAGATCAGATGCTGATAGAGTGTGAGTTGAA 1235
399 sGlyPheGlyGlnLeuHisProArgTrpGlnGlyValGluThrLeuGly 416
|||||
1236 GGGGTTTGTCATATGCATCCAGTCAGCAAGAGAGTGAACATTAGGAA 1285
416 hIleTrpSerSerSerLeuPheProAsnArgAlaProLysGlyArgVal 432
|||||
1286 CTATATACGCTCATCTATTCCTCCCAACGAGACACACCTGGAGAGGTT 1335
433 LeuLeuLeuAsnTrpIleGlyGlyAlaLysAsnProGluIleLeuSer 449
|||||
1336 CTACCTCTGAATTACATGAGAGAGCACTAATCTGCAATTTATGCAA 1385
449 sThrGlnSerGlnLeuValGluValAlaAspArgAspLeuArgLysMet 466
|||||
1386 GAGGAGACAGTGAACCTTGGAACAGTTGATCGAGATTGGAGAAATCC 1435
466 euIleLysProLysAlaGlnAspProLeuValAlaGlyValArgValTrp 482
|||||
1436 TTATTAACCCAAATGCCAGAGATCATTTGTAGTGGGGGTGAGACTGTGG 1485
483 ProGlnAlaIleProGlnPheLeuValGlyHisLeuAspThrLeuSer 499
|||||
1486 CCTCAAGCATTCACAGTCTTACTGTGCGCATCTTGATCTTCTAGATGT 1535
499 rAlaLysAlaIleMetAsnAspAsnGlyLeuGlnGlyLeuPheLeuGly 516
|||||
1536 TCGTAAAGCTTCTATCAGAAATACAGGCTTGAAGGCTCTTCTCTGGGG 1585
516 LysAsnTrpValSerGlyValAlaLeuGlyArgCysValGluGlyAlaTrp 532
|||||
1586 GTAATTATGTGTGTGTGCTGCTGAGGACATCGCTGAGAGAGCCCTAT 1635
533 GluValAlaIleSerGluValThrGlyPheLeuSerArgTrpAlaTrpLys 548
|||||
1636 GAGGATGACGCTGAGATGAACATTTCTCAAAATGAGAGTGTCAAA 1683

```

seq_name: /cgn2_6/plodata/1/ina/5B_COMB.seq:us-08-808-931-19

seq_documentation_block:

Sequence 19, Application US/08808931

Patent No. 593602

GENERAL INFORMATION:

APPLICANT: Volrath, Sandra

APPLICANT: Johnson, Marie

APPLICANT: Potter, Sharon

APPLICANT: Ward, Eric

APPLICANT: Helfetz, Peter

TITLE OF INVENTION: DNA Molecules Encoding Plant

TITLE OF INVENTION: Protoporphyrinogen Oxidase and Inhibitor-Resistant Mutants

```

TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 593602artis Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,931
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/012,705
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/013,612
FILING DATE: 28-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/020,003
FILING DATE: 21-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET INFORMATION:
TELEPHONE: (919) 541-8587
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1784 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Brassica napus (rape)
IMMEDIATE SOURCE:
CLONE: pMDC-17 (NRRL B-21615)
FEATURE:
NAME/KEY: misc feature
LOCATION: 47..1654
OTHER INFORMATION: /product= "Rape Protoc-1 coding
OTHER INFORMATION: region"
US-08-808-931-19

```

alignment_scores:

| Quality | 2006.50 | Length | 535 |
|--------------------|---------|------------------|--------|
| Ratio | 4.189 | Gaps | 4 |
| Percent Similarity | 89.533 | Percent Identity | 73.084 |

alignment_block:

US-09-508-418-2 x US-08-808-931-19 ..

Align seg 1/1 to: US-08-808-931-19 from: 1 to: 1784

```

21 ProLeuAlaPheLeuAsnArgThrSerPheIleProPheSerSerIleSe 37
|||||
68 CCGACAGCATTCCTATG.....CCATTCTCAAAATCCATT 102
37 rlysargAsnSerValAsnCysAsnGlyTrpArgThrArgCysSerVala 54
|||||
103 TCTCTGG...TCGGTCCCTACAGACCTCTCAACCTCGCTGCTCCGAT 149

```

```

54 laIysaspTyrThrValProSerAlaValAspGlyGlyProAlaIa 70
   ::::: ||||| :::::
150 CCGGTGATCGCTCGCTCTTCTCAATCGAAGCGGAGAGAGGT 199
   ::::: ||||| :::::
71 Glu.....LeuAspCysValIleValGlyAlaGlyIleSerGly 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
200 AAAACCGTCACGGCGACTCGCTCATCTCGCGCGAGAAATCAAGCCCT 249
   ::::: ||||| :::::
84 uCysIleAlaGlnValMetSerAlaAsnTyrPro.....AsnL 97
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
250 GTGATATGGCGAAGCGCTCGTAGAAGCAACCCAGCGCTGCAAAAGANTG 299
   ::::: ||||| :::::
97 eumetValThrGlnAlaArgAspArgAlaGlyGlyAsnIleThrThrVal 113
   ::::: ||||| :::::
300 TGATGGTGACGAGCGAAGACCGGTGTGGAGGAAATATCATCACCGCA 349
   ::::: ||||| :::::
114 GluArgAspGlyTyrLeuThrPglGlnGlyProAsnSerPheGlnProSe 130
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
350 GAGAGACGAGGGTTCTATGSGAAGAGGTCCCAATAGCTTTCAGCCGTC 399
   ::::: ||||| :::::
130 rAspProMetLeuThrMetAlaValAspCysGlyLeuLysAspLeuV 147
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
400 TGATCCATGCTCACTATGTGTAGATAGTGTGAAAGATGATCTAG 449
   ::::: ||||| :::::
147 alleuGlyAspProAsnAlaProArgPheValLeuTyrPylsGlyLysLeu 163
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
450 TCTTGGGAGATCTACTCTCCGAGGTTGTGTGGAAATGGGAACCTG 499
   ::::: ||||| :::::
164 ArgProValProSerLysLeuThrAspLeuProPheAspLeuMetSe 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
500 AGCCCGGTCCGTCGAACCTAAGTACGACTTGCTTCTTGACTGTAGAG 549
   ::::: ||||| :::::
180 rIleProGlyLysLeuArgAlaGlyPheGlyProIleGlyLeuArgProS 197
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
550 TATTGGAGGAGAAATTAGAGCTGGGTTGTGCCATTGGTATTCGACCTT 599
   ::::: ||||| :::::
197 erProProGlyHisGlnGlnSerValGlnGlnPheValArgArgAsnLeu 213
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
600 CACCTCCGGGTGTGAGAAATCAGTGGAAAGATTGTGAAAGCGTAACTT 649
   ::::: ||||| :::::
214 GlyGlyGlnValPheGlnArgLeuIleGlnProPheCysSerGlyValTy 230
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
650 GGTGATGAGGTTTGTGACGCGTTGATGAACCTTTCTCAGCGTGTTA 699
   ::::: ||||| :::::
230 rValAlaIysProSerLysLeuSerMetLysAlaAlaPheGlyLysValT 247
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
700 TCGGGGAGATCCTCGAAACTGATGTAAGACGCTTTTGGAAGGTTT 749
   ::::: ||||| :::::
247 rPylsLeuGlnGlnThrGlySerIleIleGlyGlyThrPheLysAla 263
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
750 GGAACCTAGAGAGAAATGTTGGAGCATCATTTGTGTCTTTAAAGCA 799
   ::::: ||||| :::::
264 IleLysGlnArgSerThrProLysAlaProArgAspProArgLeuPr 280
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
800 ATTCAAGCGAAATAAAGCTCCCAAGACAACCCGAGATCCGCGCTGCC 849
   ::::: ||||| :::::
280 oLysProLysGlnGlnThrValGlySerPheArgLysGlyLeuArgMetL 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
850 AAAGCCAAAGGGCCAAACGTTGGTTCTTTCAGAAAGAGACTCACAAATGC 899
   ::::: ||||| :::::
297 euProAspAlaIleSerAlaArgLeuGlySerLysLysLeuSerTyrP 313
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
900 TGCCAGAGGCATCTCCGAGGTTGGGTGACAAGTGAAAGTTCTCTGG 949
   ::::: ||||| :::::
314 LysLeuSerSerIleThrLysSerGlnLysGlyGlyTyrHisLeuThrTy 330
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
950 AAGCTCTCAAGTATCATAGCTGCGCAGCGAGATATACCTTAACCTTA 999
   ::::: ||||| :::::
330 rGlnThrProGlnGlnValValSerLeuGlnSerArgSerIleValMetT 347
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1000 CGAAACTCCGGAGGGATATAGTACTGTACAGACAAAGAGTGTAGTATGA 1049
   ::::: ||||| :::::
347 hrValProSerTyrValAlaSerAsnIleLeuArgProLeuSerValAla 363

```

```

1050 CTGTGCATCTCATGTGTGTAGTACTCTTGGCCCTCTCTCATCTCT 1099
   ::::: ||||| :::::
364 AlaAlaAspAlaLeuSerAsnPheTyrTyrProProValGlyAlaValTh 380
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1100 CGAGCTGAAGCCCTCTCAAACTCTACTATCCGCGAGTTGACGGCTATTC 1149
   ::::: ||||| :::::
380 rIleSerTyrProGlnGlnAlaIleArgAspGlnArgLeuValAspGlyG 397
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1150 CATTCATACGGGAAAGAAAGCAATCCGAAGCAATGCTTAATAGTGTGT 1199
   ::::: ||||| :::::
397 luleLysGlyPheGlyGlnLeuHisProArgThrGlnGlyValGlnThr 413
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1200 AACTAAAGGTTCCGCGAGTTCATCCACGACGACCAAAAGTGGAAACT 1249
   ::::: ||||| :::::
414 leuGlyThrIleTyrSerSerSerLeuPheProAsnArgAlaProLysG 430
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1250 CTTGGAACAAATATACGTTCAATCGCTCTTCCAAACCGACCGCCCTGG 1299
   ::::: ||||| :::::
430 yArgValleuLeuLeuAsnTyrIleGlyGlyAlaLysAsnProGlnIleL 447
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1300 AACAGTATTGCTATTTCATACATCGGTGAGCTACCAACACTGGGATCT 1349
   ::::: ||||| :::::
447 eusErLysThrGlnSerGlnLeuValGlnValValAspArgAspLeuArg 463
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1350 TATCAAAAGTCGAAAGGTGAGTAGTGAAGACAGTAGATAGAGACTTGAAG 1399
   ::::: ||||| :::::
464 LysMetLeuIleLysProLysAlaGlnAspProLeuValGlyValArg 480
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1400 AAGATCTGATAAAGCAAGCTTCGACCGATCCACTTGTACTTGGAGTAA 1449
   ::::: ||||| :::::
480 gValTrpProGlnAlaIleProGlnPheLeuValGlnHisLeuAspThrL 497
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1450 ATTATGGCCCAAGCCATTCCTCAGTTTCTGATAGTACATTTGATTTGG 1499
   ::::: ||||| :::::
497 eusErThrAlaLysAlaIleMetAsnAspAsnGlyLeuGlnLysLeuPhe 513
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1500 TAGACCGACGGAAGCATCGCTCTCGTCACTGTCATAGGAGGCTTATTC 1549
   ::::: ||||| :::::
514 leuGlyGlnAspTyrValSerGlyValAlaLeuGlyArgCysValGlnL 530
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1550 TTGGGTGGAAATTAATCTTCCGCTGTAGCAATTTGGTGTGTGGAAGG 1599
   ::::: ||||| :::::
530 yAlaTyrGlnValAlaSerGlnValThrGlyPheLeuSerArgTyrAlaT 547
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1600 TGCTTATGAACCTCAACCAAGTGAATGATTTTCATGCAAGTATGCTTT 1649
   ::::: ||||| :::::
547 yTyrLys 548
   |||||
1650 ACAAG 1654

```